A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

ARTICLES

Leveling and Repression 131
Philip S. Holzman and Riley W. Gardner
Effects of Meprobamate on Fear and Palmar Sweating 156
Victor G. Laties
The Effects of Diagnostic Category and Reinforcer on Learning Without Awareness 162
Allen M. Leventhal
Subliminal Effects of Verbal Stimuli 167
Gudmund J. W. Smith, Donald P. Spence, and George S. Klein
The Effect of Severity of Initiation on Liking for a Group 177
Elliot Aronson and Judson Mills
Anti-Semitism and the Displacement of Aggression 182
Leonard Berkowitz
A Test of Witkin's Field-Dependence Hypothesis 188
Hoyt H. Young
Hostility and Aversion for Expressing Hostility in Neurotics and Controls 193
H. J. Wahler
Personality Factors in Motor Learning and Reminiscence 199
Oakley S. Ray
Effects of Confidence and Motivation on Susceptibility to Informational Social Influence 204
Francis J. Di Vesta

Continued on inside front cover

Volume 59 SEPTEMBER 1959 Number 2

CONTENTS—*Continued*

Learning, Retention, and Conceptual Usage in Elderly Patients with Memory Disorder 210
James Inglis

A Comparative Study of the Personal Histories of Schizophrenic and Non-psychiatric Patients 216
William Schofield and Lucy Balian

A Circumplex Model for Maternal Behavior 226
Earl S. Schaefer

Personality Determinants of Repressive and Obsessive Reactions to Failure-Stress 236
Albert J. Caron and Michael A. Wallach

The Effects of Chlorpromazine on GSR Conditioning 246
Lonnie E. Mitchell and Melvin Zax

Selectivity in Exposure to Information 250
Judson Mills, Elliot Aronson, and Hal Robinson

The Effect of Verbal Reinforcement on the Recall of Early Memories 254
Herbert Quay

The Effects of Father-Absence on Norwegian Boys and Girls 258
David B. Lynn and William L. Sawrey

The Effect of Aversive Stimuli on Reaction Time in Schizophrenia 263
Peter J. Lang

CASE REPORT

The Elimination of Tantrum Behavior by Extinction Procedures 269
Carl D. Williams

CRITIQUE AND NOTES

Social Class and the Nature of the Incentive in Discrimination Learning 270
Glenn Terrell, Jr., Kathryn Durkin, and Melvin Wiesley

Intellectual and Personality Correlates of Test Anxiety 272
Irwin G. Sarason

Unconscious Self-Evaluation Using a Stereoscope 275
Halla Beloff and John Beloff

The Effect of Motivation Arousal on Humor Preferences 278
John F. Strickland

The Effects of Hypnotically Induced Anxiety on the Manifest Anxiety Scale and the Barron Ego-Strength Scale 281
Hanus J. Crowe and Eugene E. Levitt

Primary Process in Children as a Function of Age 284
Alvin G. Burstein

Attitudes and the Availability of Knowledge Concerning Traditional Beliefs 286
Leonard W. Doob

Subjective Probabilities of Interpersonal Relationships 290
Clinton B. De Soto and James L. Kuethe

Volume 59 SEPTEMBER, 1959 Number 2

The Journal of Abnormal and Social Psychology

Editor: M. BREWSTER SMITH, *University of California, Berkeley*

Associate Editor: E. J. SHOBEN, JR., *Teachers College, Columbia University*

CONSULTING EDITORS

GORDON W. ALLPORT
Harvard University

URIE BRONFENBRENNER
Cornell University

ROGER BROWN
Massachusetts Institute of Technology

DONALD T. CAMPBELL
Northwestern University

LAUNOR F. CARTER
System Development Corporation

RICHARD CHRISTIE
Columbia University

MORTON DEUTSCH
Bell Telephone Laboratories

I. E. FARBER
State University of Iowa

DONALD W. FISKE
University of Chicago

ROGER W. HEYNS
University of Michigan

ROBERT R. HOLT
New York University

IRVING L. JANIS
Yale University

HAROLD E. JONES
University of California

HAROLD H. KELLEY
University of Minnesota

LEONARD S. KOGAN
Community Service Society of New York

DANIEL J. LEVINSON
Harvard Medical School

DAVID C. MCCLELLAND
Harvard University

LEO POSTMAN
University of California

ELIOT H. RODNICK
Duke University

NEVITT SANFORD
University of California

SEYMOUR SARASON
Yale University

RICHARD L. SOLOMON
Harvard University

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued bimonthly in January, March, May, July, September, and November, two volumes per year. The two volumes average 832 pages. The subscription price per year is $16.00; foreign, $16.50. Single issues are $3.00.

Original manuscripts should be sent to the Editor:

M. BREWSTER SMITH
Department of Psychology
University of California
Berkeley 4, California

ARTHUR C. HOFFMAN
Managing Editor

HELEN ORR
Promotion Manager

BARBARA CUMMINGS
Editorial Assistant

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street, N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Mt. Royal and Guilford Aves., Baltimore 2, Md.
and 1333 Sixteenth Street N.W., Washington 6, D.C.

Second-class postage paid at Baltimore, Maryland, and at additional mailing offices

Copyright 1959 by the American Psychological Association, Inc.

LEVELING AND REPRESSION[1]

PHILIP S. HOLZMAN AND RILEY W. GARDNER
Menninger Foundation, Topeka, Kansas

IF A group of persons is confronted with the problem of judging the sizes of successively presented squares, each individual will perform the task in a highly idiosyncratic way that pools together his understanding of the instructions, his personal anticipations concerning the task, and his characteristic ways of dealing with situations of successive comparison in his daily life. In several publications (e.g., Gardner, 1953; Holzman, 1954; Klein, 1954), we and our colleagues have characterized the aspects of personality organization governing such individual variations in adaptive behavior as *cognitive attitudes* or *cognitive system principles*. These system principles have been described as intervening between need states (such as thirst or a drive to respond to a task) and response.

In previous papers, Klein (1954) and Holzman and Klein (1956) called attention to certain similarities between cognitive system principles and the thought modes which in psychoanalytic theory are called *defenses*. Like cognitive system principles, defenses are inferred from the patterning of cognitive functioning. For example, evidence of regression is gleaned from the use of more primitive modes of thinking, repression is inferred from the inability to recall certain ideas, and displacement is used to explain the shifting of significance from one idea or object to another. Defenses, like cognitive system principles, are mediating structures. The defenses mediate between impulses and needs on the one hand, and the pressures of reality constraints and/or internalized attitudes that conflict with impulses and needs on the other. Cognitive system principles mediate between purposes, intentions, and need states on the one hand, and adaptive requirements of a stimulus situation on the other. Our awareness of the conceptual similarity of defenses and cognitive system principles led our group to speculate that the same segment of behavior may profitably be viewed from two vantage points, in which the focus is on (*a*) its defensive implications and (*b*) its adaptive aspects. Speculation concerning relationships between the two sets of constructs employed to describe behavior from these vantage points had to await further experimental studies. The first systematic test of such relationships is described in a monograph now being prepared by Gardner, Holzman, Klein, Linton, and Spence. The monograph includes a description of results indicating that a proclivity to use repression is associated with the "leveling" extreme of the *Leveling-Sharpening* system principle and that the defense of isolation may be associated with one extreme of a system principle relevant to the extent to which individuals *scan* stimulus configurations. The present paper is concerned with an attempt to replicate the observed relationship between repression and leveling.

The cognitive system principle of *Leveling-Sharpening* refers to the degree of differentiation of memory schemata, particularly when these schemata represent successive stimulation. The lack of refinement among elements of schemata evident in test performances of subjects at the leveling extreme appears to result from relatively great assimilation of new experiences to traces (memories) of previous experiences. Thus, we assume that levelers are persons whose memory organizations contain minimally distinct individual elements. As a consequence, levelers are prone to obscure size and weight differences in tests of successive comparison.

The criterion test for the selection of levelers and sharpeners has been the Schematizing Test, described in detail by Holzman (1954). Subjects are asked to judge the sizes of 150 squares individually projected onto a screen. At the beginning of the test, the five smallest squares are presented in three different orders. The smallest square is then dropped from the

[1] The authors are indebted to David Rapaport and Gardner Murphy for critical readings of this report. The study described here was conducted as part of a larger investigation of cognitive control principles supported by a research grant (M-1182) from the National Institute of Health, Public Health Service.

series, and a square slightly but perceptibly larger than any of the first five is added. This new series of squares is then presented in three orders. In this way, the sizes of the squares gradually increase during the 35 minutes required for the test.

It has been demonstrated that while some subjects (sharpeners) keep accurate track of the general trend of increase in size over the entire test and also maintain accurate ranking of the individual squares, other subjects (levelers) tend to blur distinctions among the squares and lag behind the trend toward increasing size. Indeed, at the end of the test, many of these latter subjects were judging the 13-in. square to be 8 in. or less. Among these subjects, there was also a striking absence of differentiation between successively appearing squares, so that they at times judged a 5-in. square to be the same size as or smaller than a 3-inch square. Results of previous experiments have suggested that this process of obscuring differences and of lagging behind the trend of increase in square sizes (leveling) may be regarded as a function of the degree to which the experience of a newly appearing square "assimilates" to the aggregate of memories of preceding squares. Relationships have been demonstrated between leveling in the Schematizing Test and large assimilation effects in time-error phenomena (Holzman, 1954; Holzman & Klein, 1954), and between leveling and general constriction of ideational activity (levelers' associations remain focused around one dominant idea, which limits the occurrence of new and different themes).

The similarity between leveling and repression is detailed in the monograph by Gardner, Holzman, Klein, Linton, and Spence:[2]

> In psychoanalytic writings, repression proper, as distinguished from primal repression (Freud, 1915), is conceived of as the expulsion of an idea from consciousness in order to avoid experiencing the associated unpleasant affect, usually anxiety. According to Freud, such expulsion of a conscious idea is in part determined by its attraction to an idea that is already unconscious . . . . Thus, the process of repression is at least analogous to the process of assimilation apparently measured in our laboratory tests of *Leveling-Sharpening*. In both repression and leveling, a new idea or experience assimilates to an older one by virtue of the relationship between them. As a consequence, the assimilated idea loses its individuality, and hence its availability to consciousness as a discrete entity. It seems plausible to assume that repression as a typical way of coping with conflict can occur most easily in persons whose memory organization is congenial to this kind of easy assimilation. Where such a process is not a natural outcome of a person's basic patterns of memory organization, different ways of dealing with traumatic material may appear.

[2] David Rapaport points out that this formulation of the process of repression is an oversimplification. It may be more profitable to regard the process as one in which the preconscious idea becomes invested with drive cathexis and thus becomes liable to repression.

The results obtained earlier indicated that extremely repressive persons tend to be levelers, although not all levelers are extremely repressive. It was hypothesized that the same relationships would appear in the present study.

METHOD

Subjects

The *S*s were 10 extreme levelers and 10 extreme sharpeners, selected from a larger group of 80 by means of their performances in the Schematizing Test. They were female university students and ranged in age from 18 to 21.

Procedure

The Schematizing Test was administered in groups of two to five *S*s.

In the present experiment, a double criterion involving both ranking accuracy and degree of lag was used to select extreme levelers and sharpeners. The ranking accuracy score pertains to the degree to which *S*s accurately rank the sizes of the squares, that is, the extent to which *S*s recognize size differences between stimuli. We have assumed that a low ranking accuracy score reflects a retention process that does not preserve distinctions in the incoming stimuli. We have also assumed that a high degree of lag reflects a tendency for impressions of stimuli to be highly influenced by anticipations about the series or by impressions built up about the series, a process that also obscures the discreteness of presently appearing stimuli. Both low ranking accuracy and high lag are presumed to be functions of relatively great assimilation between a presently appearing stimulus and the trace residue of preceding stimuli.

The Rorschach test was administered individually to the 10 extreme levelers and 10 extreme sharpeners by an experienced tester. The 20 records were then rated on a four-point scale indicating the degree of the *S*'s reliance on repression as a defense. There were four raters, each of whom had had extensive experience in interpreting psychological tests at the Menninger Foundation. None of the raters knew which *S*s were levelers or sharpeners, and two of the raters knew nothing of the purpose of the experiment.

Each rater was provided with a set of specific criteria as a guide for inferring repression from Rorschach protocols. These criteria, based upon rationales worked out by Rapaport, Gill, and Schafer (1946), included the following as indicators of reliance upon repression: few responses; few movement responses; emphasis upon color; stereotypy of content; failure to respond to one or more cards; child-like percepts or verbalizations; weak integration of the inkblot material; vagueness; "affective" or phobic comments about the inkblots; and the reporting of symbolic percepts without awareness of their significance. Not all of these criteria, of course, needed to be present in the records for repression to be inferred.

RESULTS

1. The reliability[3] of the average ratings for the four raters was .79, indicating high consistency in judgments of the degree to which *S*s relied upon repression.

2. Six *S*s received an average rating of 3.00 or above, indicating a strong tendency to use repression as their principal defense. All of these six extreme repressers were also extreme levelers. Such a relationship could occur by chance less than two times in 100 (the binomial probability is .0156 for a one-tailed test). The confidence one may place in this relationship is increased by the fact that it confirms the results of a previous study. As in the preceding study, however, the converse is not true; not all extreme levelers rely upon repression as a principal defense. Of the remaining 14 *S*s, who were rated lower in reliance upon repression, 10 were sharpeners and four levelers.

DISCUSSION

Up to the last decade of the 19th century, neurotic phenomena, and hysteria in particular, were viewed as outcomes of hereditary dispositions. The first systematic use of the concept of defense appeared in Freud's early attempts (1894, republished 1924) to explain the appearance of certain neurotic symptoms. Freud's explanation invoked the idea of a dynamic struggle between two incompatible ideas as the precondition of a neurosis. An idea, representing a wish, in itself pleasurable, conflicts with a set of dominant attitudes (Freud's early ego concept), and a struggle ensues. If the ego attitudes are powerful enough, they are able to prevent awareness of the contradictory wish. Freud called this process of banishment *defense*. Thus, in Freud's early conception, defense referred to a purpose: that of maintaining the integrity of the "dominant ideational complex."

In his subsequent work, Freud elaborated and expanded the concept defense. Energy dynamics were added, and the concept was further enriched by elaboration of the nature of the unconscious and development of a more systematic view of the ego as a structure. From his first bold, yet elegantly simple, description of repression as the excluding from consciousness of the memory of an actual experience, Freud (1925) developed the complex theory of primal repression and secondary expulsion. The phenomonenon he sought to explain remained, "How is it that patients do not remember dynamically important ideas, events, fantasies?"

In the work subsequent to 1926 on the psychology of ego processes, much attention was paid to the patterning of defenses, the precipitants of defenses, and detailed description of defenses. In these efforts, however, the essential *processes* by which the ego accomplishes its defensive purposes were ignored. Waelder (1951), for example, attempted to analyze projection into its component parts. One can, of course, analyze concepts into their components; one can make explicit what is implicit. But in doing this, one must keep in mind that he is dealing with conceptual material and not empirical material. To break up projection into component defenses tells us that projection is not a simple construct. But whether such efforts tell us something new about the *process* of projection is open to question.

What may be needed is a searching perusal of the *processes* that the defense names refer to and of relations between these and other ego processes. The question becomes not "What defenses made up reaction formation?" but "How is it possible for a person to convert hate for another person into a overweening politeness toward the same person, and at the same time be unaware of his hate?" In respect to repression, we would ask, "By what manipulations within the mind is it possible to forget the significant idea?"

[3] This reliability coefficient was the intraclass r for four raters (Ebel, 1951).

The writings of Anna Freud (1937) and Hartmann (1939, republished 1958) contain significant exceptions to the generally narrow approach to the study of defensive processes. They saw clearly that the defensive activities of the ego are not a separate and distinct category of response. Hartmann (1939, republished 1958), for example, wrote: "I want to stress only the point I have made repeatedly, the defense processes may *simultaneously* serve both the control of instinctual drive and adaptation to the external world. . . . We observe at every turn that the ego simultaneously effects adaptation, inhibition, and synthesis (p. 51)." He added (p. 106) that "We may also ask whether or not the defense processes are influenced by the maturation and exercise of the apparatuses of the conflict-free ego sphere."

Our results likewise suggest a *linking* of adaptive and defensive processes. The exact nature of the link, however, is not determined. Four possible links present themselves: (*a*) Repression and leveling are identical processes viewed from different vantage points. (*b*) Repression and leveling are two different processes, but both differentiate from a common matrix and therefore bear the traces of their common source. (*c*) Leveling is a consequence of the repressive solution of conflicts. (*d*) Repression relies in part, but by no means exclusively, upon the process of leveling for its execution. This fourth possibility is the explanation we favor. That is, defensive forgetting makes use of the memory apparatuses at the ego's disposal and takes its form from the major characteristics of these apparatuses. Both defense and adaptation have been conceptualized as functions of the ego. Since the ego executes its functions via the available psychic apparatuses, the major characteristics of these apparatuses will be apparent in both defensive and adaptive activities. Thus, if the tendency toward extreme assimilation among the mental representatives of successively appearing stimuli is a relatively stable characteristic of a particular cognitive structure, this quality may be apparent not only when the apparatus is used adaptively but when it is used defensively as well. Knowing that a subject relies chiefly upon repression for defensive purposes, then, allows us to infer that leveling is a prominent mode of cognitive functioning. The converse, however, does not appear to be true. That is, a tendency toward extreme assimilation of successively appearing stimuli does not *insure* the use of repression for defense. We would, however, offer the hypothesis that other cognitive system principles provide the form for the favored defensive maneuvers; that a necessary but still insufficient condition for choice of a particular defense is the characteristics of the ego apparatuses involved in the defensive operations, and that an investigation of defense must include a study of general cognitive functioning. Our results suggest that repression, as a principal defense, occurs in a cognitive structure that favors a general assimilative or leveling attitude toward stimuli.

SUMMARY

Subjects relying chiefly on repression as a defense mechanism were shown to be levelers in a neutral psychophysical task. This finding replicates a previous finding (Gardner et al., in press). The statistical linkage of defense and cognitive system principle suggests that the formal properties of cognitive structures provide the occasion for the choice of defense.

REFERENCES

EBEL, R. L. Estimation of the reliability or ratings. *Psychometrika*, 1951, **16**, 431–437.

FREUD, A. *The ego and the mechanisms of defence.* London: Hogarth, 1937.

FREUD, S. The defence neuro-psychoses. (Originally published 1894) *Collected papers.* Vol. I. London: Hogarth, 1924. Pp. 59–75.

FREUD, S. Repression. (Originally published 1915) *Collected Papers.* Vol. IV. London: Hogarth, 1925. Pp. 84–97.

GARDNER, R. W. Cognitive styles in categorizing behavior. *J. Pers.*, 1953, **22**, 214–233.

HARTMANN, H. *Ego psychology and the problem of adaptation.* (D. Rapaport, Trans.) (Originally published 1939) New York: Int. Universities Press, 1958.

HOLZMAN, P. S. The relation of assimilation tendencies in visual, auditory, and kinesthetic time-error to cognitive attitudes of leveling and sharpening. *J. Pers.*, 1954, **22**, 375–394.

HOLZMAN, P. S., & KLEIN, G. S. Cognitive system principles of leveling and sharpening: Individual

differences in assimilation effects in visual time error. *J. Psychol.*, 1954, **37**, 105–122.

HOLZMAN, P. S., & KLEIN, G. S. Motive and style in reality contact. *Bull. Menninger Clin.*, 1956, **20**, 181–191.

KLEIN, G. S. Need and regulation. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln: Univer. Nebraska Press, 1954.

RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. II. Chicago: Yearbook Publishers, 1946.

WAELDER, R. The structure of paranoid ideas: A critical survey of various theories. *Int. J. Psychoanal.*, 1951, **32**, 167–177.

Received April 7, 1959. Prior publication.

EFFECTS OF MEPROBAMATE ON FEAR AND PALMAR SWEATING[1]

VICTOR G. LATIES

School of Medicine, Johns Hopkins University

THERE are reports that meprobamate (Miltown, Equanil) can decrease the magnitude of fear or "anxiety" responses in man (Laties & Weiss, 1958). The purpose of this study was to investigate the effects of meprobamate on several such responses of normal human subjects who were placed in a mild stress situation. Palmar sweating was the main response measured. Several investigators have shown it to be elevated in stress situations (Davis, 1957; McCleary, 1953). Data were also gathered on several self report measures. The mild stress was a ride on a ferris wheel.

METHOD

Twenty paid volunteers, 10 male and 10 female, ranging in age from 19 to 30 (median 23.5) were each studied twice. Five of each sex took a total of 1600 mg. of meprobamate divided into four equal doses, at 8 a.m., noon, 4 p.m. and 6 p.m., on their first day, and identical appearing placebos on their second day. The other 10 subjects (*S*s) took the placebo first and the drug second. The *S*s, almost all medical students or technicians, were unaware of the drug under study. They were also told nothing about a placebo's being involved.

After taking the fourth tablet on a given experimental day, *S*s went to the carnival grounds where the apparatus used to induce mild fear was located. It was a 45-ft. high, 12-seat ferris wheel, known as the Big Eli Wheel No. 12 (Eli Bridge Co., Jacksonville, Ill.), which revolved about six times per minute. Thus *S*s travelled at approximately 14 ft. per sec. An initial measure of palmar sweating was obtained during a 15-min. period shortly after arrival at the carnival. A second 15-min. determination was started approximately 45 min. later and included a 10-min. ride on the ferris wheel. *S*s rode one to a seat. Each time the wheel revolved, they made self ratings of the amount of fear experienced by writing on a card a number from 1 to 10. The numbers corresponded to those on a "fear thermometer" printed on the cards (Walk, 1956). No. 1 was labelled "completely relaxed"; No. 10 was labelled "as scared as I've ever been." After the ride, *S*s used the same scale to rate retrospectively the amount of fear experienced at three times during the experiment: during the pre-ride determination, just before they boarded the wheel, and during the first two or three loops. At this time they were also asked to report whether they had experienced any of the following symptoms while on the wheel: increased heart rate, trembling hands, upset stomach, cold sweat, and nervousness. Space was also allowed on the form to report any other symptoms experienced.

The USAF Manual Perspiration Test developed by McCleary (1953) was used to measure palmar sweating. Anhydrous silica gel (5.2 gm.) was put into a 3 x 1¾-in. cotton marquisette bag. The open end was clipped shut and the bag placed in a 40 x 75 mm. weighing bottle with a standard taper ground joint. The leakage rate was determined for each bottle (0.9–.16 mg. per hour) and subsequent measurements corrected accordingly. For each sample, the *S* took the bag out of the bottle, clenched it between his palm and the palmar aspects of his fingers for precisely 15 min., and then returned it to the bottle. The change in weight of the bottle and its contents following this procedure was taken as the measure of palmar sweat. Weighings were made on an analytical balance to the nearest milligram. In order to check on whether mere travel on the ferris wheel led to differential absorption of moisture, four other *S*s were run through the procedure carrying the sweat bags in a rubber-gloved hand. Since it was discovered that the bags picked up a mean of 12 mg. more moisture when carried on the ferris wheel than they did when held by the same *S*s for a like period on the ground, all determinations made for the period including the ride were reduced by this amount. The mean temperature during placebo runs was 28.5°C, during meprobamate, 28.1°C.

After analysis of the data from this main experiment, a supplementary study was executed in the laboratory to obtain more data on the palmar sweat response to the drug. Sixteen more paid volunteers, 6 male and 10 female, ranging in age from 21 to 38 (median 24.5), held bags of anhydrous silica gel for 15-min. periods on two nonsuccessive days. Prior to each period, *S*s ingested either a total of 1600 mg. meprobamate or placebos. These experiments were all run at 3 or 4 p.m. Equal doses of drug (or placebo) were ingested before retiring the previous night, and one, five, and nine hours prior to the sweat determination. Ten *S*s took placebos before their first session and the drug before their second; six *S*s ran in the reverse order. Mean temperature was 25.5°C for both placebo and meprobamate runs.

[1] This research was supported in part by a grant from the United States Public Health Service, National Institutes of Health, No. B-865-C, and by a grant from the Burroughs Wellcome & Company. The author wishes to thank Donald R. Goodenough and Paul Meier for their suggestions concerning the statistical treatment of the data. Frank M. Berger, Wallace Laboratories, New Brunswick, N. J., kindly supplied the meprobamate.

RESULTS

Palmar sweating. The ferris wheel ride led to a mean increase in palmar sweating of 102.7 mg.: from 259.8 mg. for the pre-ride period to 362.5 mg. for the period including the ride. An analysis of variance indicated that this difference was significant at the .01 level (Lindquist, 1953). The drug versus placebo comparison did not yield such a clear-cut result. The mean amount sweated was 335.1 mg. under placebo, and 287.1 mg. under meprobamate. This difference of 48.0 mg. yielded an F of 3.90, but an F of 4.49 is required for significance at the .05 level (for 1 and 16 degrees of freedom).

The supplementary experiment was executed to test whether the drug variable would yield a significant difference when the environment was held more constant than possible in the field. The mean amount of palmar sweat was 288.1 mg. for *S*s with placebo, and 239.8 mg. for *S*s with meprobamate. The difference of 48.3 mg. was almost precisely the same as that found in the main experiment. A t test for paired observations (Edwards, 1954) was significant at the .02 level. (All tests of significance cited are two-tailed.)

From the two experiments, it appears that meprobamate, in the dose used, decreased palmar sweating. The interaction between the ferris wheel stress and drug variables in the main experiment was not significant, with an F value less than one, i.e., the ferris wheel stress failed to produce a significantly smaller

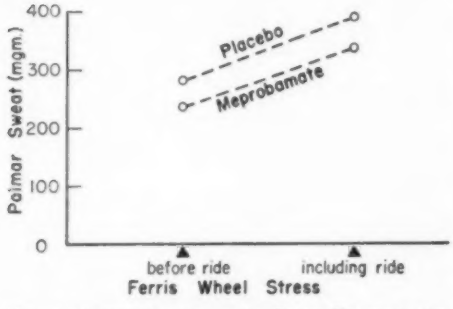

FIG. 1. PALMAR SWEATING OF *S*s ON PLACEBO AND MEPROBAMATE DURING 15-MIN. PERIODS BEFORE AND INCLUDING THE FERRIS WHEEL RIDE

TABLE 1
WEIGHTS FOR INDEX OF UNPLEASANT SYMPTOMS

Symptom	Assigned Weight
Upset stomach	3.6
Upset stomach (slight)	2.0
Nervousness	2.2
Nervousness (slight)	1.2
Increased heart rate	2.0
Increased heart rate (slight)	1.0
Hand trembling	2.2
Hand trembling (slight)	1.2
Dizziness	2.2
Dizziness (slight)	1.0
Cold sweat	2.6

increment in palmar sweating when the *S*s were on meprobamate than when the same *S*s were on placebo (see Fig. 1).

Unpleasant symptoms. An index was constructed from the reports made by *S*s about their unpleasant symptoms during the ride by having five professional people in the laboratory (three MDs, two PhDs) rate the symptoms in terms of their "severity" from one to four, with four being the most unpleasant. Table 1 shows the mean weights given by the judges and used in the analysis. The mean total scores on this index were 3.44 for placebo, and 2.12 for meprobamate. This decrease was significant at the .01 level. There is no good estimate of what *S*s would have reported on the ground before the ride, but it would appear that they responded to the ride with fewer unpleasant symptoms when under meprobamate.

Concurrent fear ratings. Figure 2 shows the mean fear ratings made by *S*s while riding the ferris wheel. The curve shows that the ride produced a mild amount of self-rated fear. The highest amount occurred on the initial revolutions, with means of 4.55 for placebo and 4.30 for meprobamate. The response extinguishes rapidly, and the curves become essentially asymptotic between the fifteenth and twentieth turns. The obviously significant decrease is evidence that the ferris wheel ride did indeed produce a mild amount of fear. All *S*s showed this drop.

The curves for placebo and meprobamate are indistinguishable. Analyses of variance were carried out on the sums of the first three

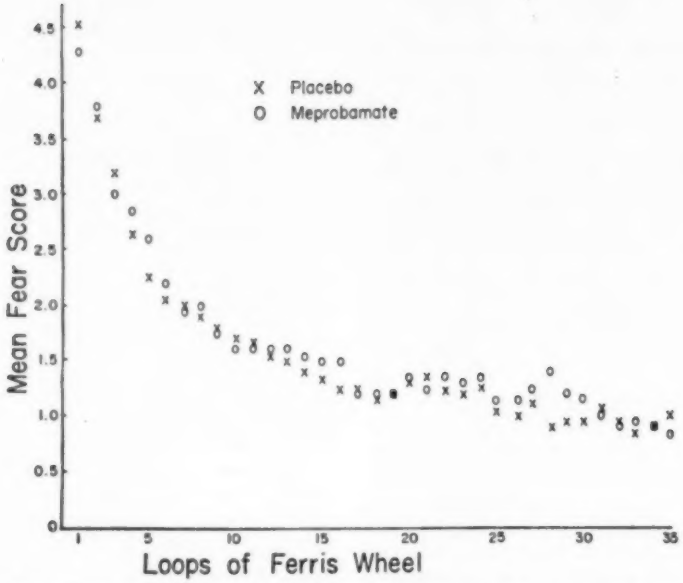

FIG. 2. SELF-RATINGS OF FEAR MADE BY *Ss* ON PLACEBO AND MEPROBAMATE DURING THE FERRIS WHEEL RIDE

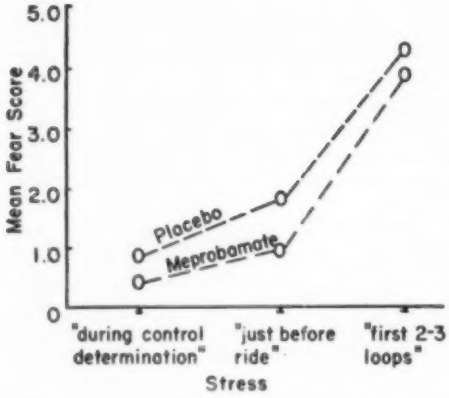

FIG. 3. SELF-RATINGS OF FEAR MADE BY *Ss* ON PLACEBO AND MEPROBAMATE AFTER THE FERRIS WHEEL RIDE

ratings and on the sums of the second three ratings. An F of less than one resulted each time for the drug variable. If one assumes that stress decreased as the ferris wheel ride continued, then the complete lack of separation between the curves for placebo and meprobamate argues for the absence of any interaction between the effect of stress and the drug state of *S*.

Retrospective fear ratings. The three estimates of fear made after the ride are graphed in Fig. 3. The difference between "during control determination" and "just before ride" is statistically significant at the .01 level by a simple sign test (Dixon & Massey, 1951); so are the differences between either of the pre-ride estimates and the "first 2-3 loops" rating. By t tests for paired observations, there was no evidence that the drug and placebo curves differed, whether the three estimates were considered separately or pooled.

Relaxation and sleepiness. Following the ferris wheel ride, *Ss* were asked via a questionnaire, "Did you feel different in any way today? (More tense? More relaxed? Happier? Sadder? etc.)." Responses other than ones concerned with relaxation and sleepiness did not occur with a frequency high enough to be meaningfully analyzed. For the same reason,

reports of sleepiness were pooled with those of relaxation. When each *S*'s replies under placebo and meprobamate were compared, 6 *S*s reported more relaxation or sleepiness under meprobamate than under placebo, 1 reported more under placebo, and 13 reported no difference. In the supplementary laboratory study, *S*s were orally quizzed as to their relaxation and sleepiness during each of the two days of the experiment. Nine of the 16 *S*s reported having been more relaxed or sleepy under the drug than under the placebo, 2 reported more under the placebo, and 5 reported no difference. When the results of the two experiments are combined, 18 of the total of 36 *S*s noticed no difference from one treatment to the other, 15 were more relaxed or sleepy under meprobamate, and 3 were more so under placebo. This difference is significant at the .01 level by a sign test.

Both palmar sweating and concurrent fear ratings were studied for *S*s who experienced daytime sedation from the drug and likewise for those who could detect no such effect. Neither measure gave any indication that the sedative effect of the drug was related to its other effects.

Session effects. Since each *S* was studied twice, it was possible to examine some of the effects of repeated exposure to similar stimulus conditions. The amount sweated during the pre-ride period did not change. It was 260.9 mg. for Session I, 258.6 mg. for Session II. The means for the period including the ferris wheel ride were 339.9 mg. and 385.1 mg. for Session I and Session II respectively. Thus, the ride led to significantly more palmar sweating ($p < .01$) during the second session than during the first session.

The index of unpleasant symptoms experienced during the ride showed no significant change with sessions. The mean total for Session I was 2.89, for Session II, 2.67.

Experience with the wheel led to a decrease in its fear producing properties if fear is measured by the reports of the *S*s. On the concurrent fear ratings, the mean rating during the first three revolutions of the wheel dropped from 4.32 for Session I to 3.20 for Session II. This difference is statistically significant at the .01 level. Analysis of the totals for the second three loops yielded a similarly significant difference between the means of 3.02 for Session I and 1.82 for Session II. Retrospective fear ratings also showed a decline from the first to the second session. The ratings that referred to the pre-ride period were compared to the ratings that referred to the first two or three loops of the ride. The mean change was from 4.11 for Session I to 2.93 for Session II. This difference is significant at the .05 level.

DISCUSSION

It seems clear that meprobamate affected the measure that reflects predominantly autonomic activity while leaving relatively untouched the self ratings of fear. Such a difference cannot be due to a total lack of sensitivity of the fear ratings since they were changed by the ferris wheel ride. It may be that a relative difference in sensitivity of the measures is at issue here. Thus, the larger (approximately 100 mg.) change in palmar sweating induced by the ferris wheel ride was accompanied by a significant change on self ratings; the smaller (approximately 50 mg.) change induced by the drug was not.

Of perhaps greater interest than the simple effects of meprobamate on the measures indicative of fear is the relative lack of effect of the drug on these same measures when the response to the ferris wheel ride is viewed as the dependent variable. This effect is given by the interaction of the drug and stress variables. The drug led to a mean decrease of 44.1 mg. in palmar sweating when the pre-ride periods are compared and a mean decrease of 51.8 mg. when periods including the ferris wheel ride are similarly compared (Fig. 1). This difference in change is not statistically significant. Thus the drug, while decreasing palmar sweating, did so in a way that appears not to be related to stress level. This may mean that the action of the drug is directly upon the sweat mechanism.

The same lack of correspondence with stress level appears for the self ratings with no discernible difference induced by the drug at either stress level. Said in another way, the stress produced almost equal increases of 98.8 mg. when *S*s were on the drug and 106.6 mg. when they were on placebo. It was only with

the relatively gross measure of reported unpleasant symptoms that the drug seemed to have affected the response to the mild stress situation.

The drug-induced decrease in palmar sweating stands in direct contradiction to results reported recently by Marquis, Kelly, Miller, Gerard, and Rapoport (1957). They measured palmar sweating with the thumb print technique reported by Mowrer (1953). Readings taken during a mock driving task were unexpectedly higher after 800 mg. of meprobamate than under placebo.

The significance of this finding is unclear, since the effect did not reappear upon replication of the study.[2] In their second study, measurements of palmar sweating were also made before the driving task, thereby allowing for a test of whether the changes due to the stress associated with the task differed for placebo and drug. They did not.

A seeming discrepancy arises when the *decrease* in verbal reports of fear from the first to second session is contrasted with the *increase* in palmar sweating reaction to the stress. A somewhat similar finding was reported by Blatz (1925). He found that naive *S*s reported fear and showed heart rate changes after being tipped over unexpectedly in a trick chair. When *S*s were given the same treatment on another occasion, the heart rate changes persisted, although to a lesser degree, but this time Blatz's *S*s did not report fear. A verbal report of fear may extinguish relatively rapidly when compared to an autonomic indicator. But the autonomic indicator used here not only did not decrease significantly; it actually increased! The results can be tentatively accounted for if one assumes that the sweat response became conditioned to the stimuli associated with the ride during the first session. The sweat response at the second session would thus be a composite reaction to those newly conditioned stimuli and to the stimulation that was originally effective in producing sweating. While the response to the original stimuli could have decreased, the addition of the new stimuli may have more than compensated for this decrement. The data do not afford a direct test of the hypothesis, but one can test a derivation from this formulation.

If the postulated conditioning occurs, it should lead to the largest response increment for the *S*s who are most discomfitted by the ride. Conversely, *S*s who go through their first experience with the ferris wheel with relatively minor reactions should not carry into their second experience much in the way of newly conditioned responses to stimuli connected to the ride. This contention was tested by examining the relationship of the palmar sweating response to scores on the unpleasant symptom index during the first session. The amount sweated during the pre-ride period was subtracted from that sweated during the period which included the ride. The difference from Session I was subtracted from the difference from Session II to yield the increment in sweating for the *S*'s second experience with the ferris wheel over his first. These increments were then correlated with scores on the index of unpleasant symptoms from Session I, separately for each of the two drug administration orders. These two Pearsonian correlations were combined via a Z-transformation (Edwards, 1954). The resulting correlation, although only .47, is significant at the .05 level for 18 degrees of freedom. The suspicion that the increase in palmar sweating response to the second ferris wheel ride may have been due to conditioning to stimuli associated with the first ride thus has some indirect support.

[2] Uhr, L. Personal communication. August 26, 1957.

SUMMARY

A short ride on a ferris wheel was found to yield moderate self ratings of fear, to increase palmar sweating, and to lead to reports of unpleasant somatic symptoms. A 1600-mg. divided dose of meprobamate, as compared to a placebo, produced a reduction in palmar sweating both before and during the ride and a reduction in reported unpleasant symptoms during the ride, but led to no detectable changes in self ratings of fear.

Except for the index of unpleasant symptoms experienced during the ride, no evidence was found that the ferris wheel stress affected subjects any differently when they had taken meprobamate than when they had taken placebo.

Subjects sweated significantly *more* during the second ferris wheel ride than during the first, but reported significantly *less* fear during their second ride. A possible interpretation was offered in terms of (*a*) verbal fear responses extinguishing more rapidly than autonomic responses and (*b*) simple conditioning of the autonomic response (palmar sweating) to stimuli associated with the first ferris wheel ride.

REFERENCES

BLATZ, W. The cardiac, respiratory, and electrical phenomena involved in the emotion of fear. *J. exp. Psychol.*, 1925, **8**, 109–132.

DAVIS, R. H. A further study of the effect of stress on palmar prints. *J. abnorm. soc. Psychol.*, 1957, **55**, 132.

DIXON, W. J., & MASSEY, F. J., JR. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.

EDWARDS, A. L. *Statistical methods for the behavioral sciences.* New York: Rinehart, 1954.

LATIES, V. G., & WEISS, B. A critical review of the efficacy of meprobamate (Miltown, Equanil) in treatment of anxiety. *J. chron. Dis.*, 1958, **7**, 500–519.

LINDQUIST, E. L. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

MCCLEARY, R. A. A simple method for the physiological measurement of anxiety. *J. aviat. Med.*, 1953, **24**, 508–513; 535.

MCCLEARY, R. A. Palmar sweat as an index of anxiety: A field method suitable for large groups. *USAF Sch. Aviat. Med. Rep.*, 1953, Proj. No. 21-1207-0004, Report No. 1.

MARQUIS, D. G., KELLY, E. L., MILLER, J. G., GERARD, R. W., & RAPOPORT, A. Experimental studies of behavioral effects of meprobamate on normal subjects. *Ann. NY Acad. Sci.*, 1957, **67**, 701–710.

MOWRER, O. H. *Psychotherapy: Theory and research.* New York: Ronald, 1953.

WALK, R. Self ratings of fear in a fear-invoking situation. *J. abnorm. soc. Psychol.*, 1956, **52**, 171–178.

Received January 24, 1958.

THE EFFECTS OF DIAGNOSTIC CATEGORY AND REINFORCER ON LEARNING WITHOUT AWARENESS[1]

ALLAN M. LEVENTHAL

State University of Iowa

IN RECENT years there has been a notable increase in the experimental investigation of learning without awareness. Much of the stimulus for this work was furnished by a study by Greenspoon (1955). This investigator was able to show in an "interpersonal" experimental situation that modifications in *S*'s behavior contingent on that of *E* could be accomplished without *S*'s awareness. Similarly, as shown by Adams's (1957) review, many other studies have reported effects of the verbal behavior of *E* on the verbal behavior of *S*s. These investigations offer good evidence that verbal stimuli can operate as reinforcers in an experimental situation and produce conditioning phenomena in both normal and psychiatric populations. In general, *E*s have varied characteristics of the subject population or properties of the reinforcing agent. The present investigation attempted to vary systematically aspects of *both* variables, utilizing *S*s discriminated on the basis of diagnosis and reinforcers involving approval or disapproval.

METHOD

Subjects

The *S*s were 180 white males diagnostically subdivided in the following manner: 60 schizophrenics, 60 nonpsychotic psychiatric patients (psychoneurotics or personality disorders), and 60 ambulatory patients from the general medical hospital population (excluding neurology and psychiatry) who were presumably "normal." The patients were drawn from six hospitals in Iowa. In addition to diagnosis, patients were selected on the basis of their being between 20 and 49 years of age, without evidence of brain injury, hospitalized for the first time (excepting any military hospitalization) and for less than four months, and literate.

Procedure

The materials for the experiment consisted of 80 3 x 5 white index cards. Eighty commonly used and neutrally toned verbs were selected, and a different verb was typed in the middle of each card. All verbs were in the past tense. Four pronouns—I, We, He, They—were placed below each verb with their respective orders randomized for the series of 80 cards.

Upon entering the experimental room, each *S* was seated before *E* at a table and told he had been called to participate in some hospital research. When *S* agreed to participate, he was instructed to make up a sentence containing the verb indicated on the card and beginning with any one of the pronouns.

In order to establish a base level, all *S*s were treated alike during the first 20 card presentations. *E* exposed each stimulus card, recorded *S*'s sentence, and then exposed the next card. This sequence of 20 card presentations constituted Block 1. The following 60 presentations were similarly divided into Blocks 2, 3, and 4. All cards were thoroughly shuffled before each *S* was tested.

Within each diagnostic sample, each *S* was randomly assigned to one of four treatment groups defined in terms of *E*'s behavior over Blocks 2–4. These groups included:

Group R, a *partial reward* group. *E* said "good" immediately following any sentence starting with the word I or We. *E* made no response following any sentence starting with the word He or They.

Group P, a partial *punishment* group. *E* said "not so good" immediately following any sentence starting with the word He or They, but made no response following any sentence beginning with the word I or We.

Group RP, a continuously *rewarded and punished* group. *E* said "good" immediately following any sentence starting with the word I or We and "not so good" immediately following any sentence beginning with the word He or They.

Group C, a *control* group. *E* made no response following any sentence.

In order to test awareness of the relation between *S*'s responses and those of *E*, each *S* was asked the following questions after he had completed the eighty trials: (*a*) What did you think the experiment was about? (*b*) Do you think you used any of the pronouns any more than any of the others? If so, why? (*c*) Is there anything I did that might have influenced the kind of sentences you made up? (*d*) Do you remember my saying "good" (or "not so good") every so often? (*e*) What did you think it meant?

Elimination of Subjects

*S*s were eliminated for using any one or two pronouns for any 12 consecutive sentences during the first 20 card presentations (five normals, seven neurotics, and two schizophrenics were so eliminated) and for displaying awareness of the purpose of the experiment during the question period (seven normals, five neu-

[1] Submitted in partial fulfillment of the requirements for the Ph.D. degree in psychology at the State University of Iowa. Gratitude is extended to Leonard D. Goodstein for his assistance in the preparation of the dissertation on which this paper is based.

TABLE 1
ANALYSIS OF VARIANCE FOR ALL GROUPS FOR BLOCKS ONE THROUGH FOUR

Source	*df*	*MS*	*F*	*p*
Treatments (T)	3	297.38	10.16	<.001
Diagnoses (D)	2	15.88	<1.00	—
T × D	6	33.75	1.15	—
error (b)	168	29.27		
Trials (t)	3	298.97	51.37	<.001
t × T	9	46.57	8.00	<.001
t × D	6	3.97	<1.00	—
t × T × D	18	41.69	7.16	<.001
error (w)	504	5.82		

rotics, and two schizophrenics were so eliminated). Position preferences were warned against if they occurred in the first block of trials, and no *S* had to be eliminated for persisting in this tendency.

RESULTS

None of the differences among the means for age or education of the diagnostic groups proved to be statistically significant. Therefore, any group differences in learning ability cannot be accounted for simply on the basis of educational level (a gross measure of intellectual ability) or age.

The frequency of sentences starting with I or We within each block of trials was taken as the criterion measure for all *S*s. As the *S*s had to respond with one of the two groups of pronouns, the effect of positive reinforcement of the I and We class and of negative reinforcement of the He and They class was shown by an increase in usage of the I and We class. The degree to which *S*s used I and We to begin their first 20 sentences was chosen as a measure of the basal level of usage of the first person on this task. To determine whether significant differences existed among the diagnostic groups or the planned treatment groups at the outset of training, Block 1 data were analyzed by a factorial design (Lindquist, 1953). Results of this test were nonsignificant.

Graphic depiction of the results for each of the treatment conditions within each diagnostic category is presented in Figures 1, 2, and 3. Analysis of these data was by a Type III "mixed" design (Lindquist, 1953), in which diagnoses and treatments constituted "between *S*s" factors, whereas trials and those interactive effects which included trials were "within *S*s" factors. The summary

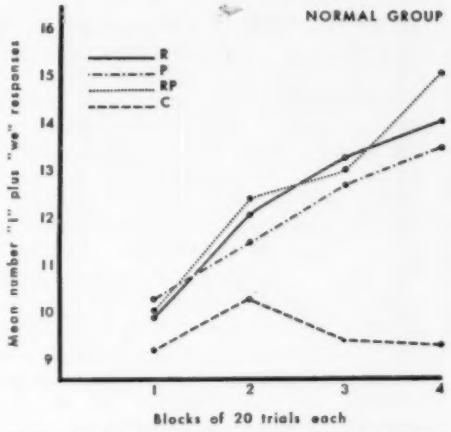

FIG. 1. MEAN NUMBER OF "I" PLUS "WE" RESPONSES FOR THE NORMAL GROUP OVER THE FOUR BLOCKS OF 20 TRIALS EACH

table of the results of this analysis of variance is presented in Table 1.

Table 1 indicates that the triple interaction of trials by treatments by diagnoses was highly significant ($p < .001$), indicating that the trials by treatments interaction effects vary according to diagnosis, thereby necessitating a direct study of the simple effects rather than an interpretation of the lower-order interactions, the main effects, or both. Accordingly, Type I incomplete factorial designs (Lindquist, 1953) were employed to test the significance of the simple effects.

Normal group. Analysis of the data for the normal *S*s (Fig. 1) disclosed that the trials by treatments interaction effect was significant ($p < .025$). If it is assumed that the various treatments were cumulative in their effects, then the most satisfactory criterion measure of the treatment effects would be the means on the last block of trials, when such effects would be maximized. Accordingly, the critical difference (Lindquist, 1953) among the means necessary for significance at the .05 level was calculated. Results indicated that Groups *R*, *P*, and *RP* all differed significantly from Group *C*, but not from each other.

Neurotic group. The results for the neurotic *S*s, presented graphically in Fig. 2, indicate the presence of a significant trials by treatments interaction ($p < .001$). Again, the

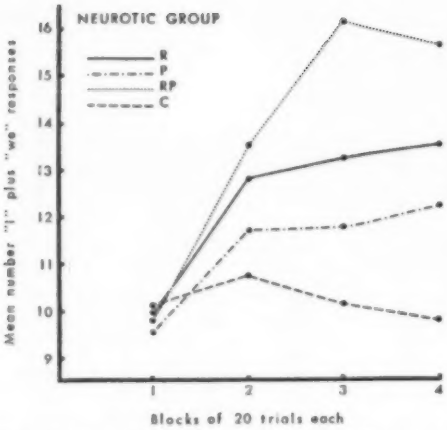

FIG. 2. MEAN NUMBER OF "I" PLUS "WE" RESPONSES FOR THE NEUROTIC GROUP OVER THE FOUR BLOCKS OF 20 TRIALS EACH

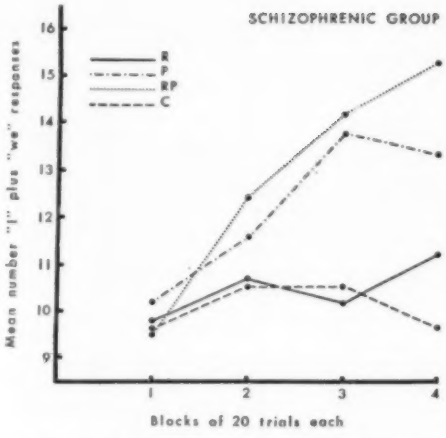

FIG. 3. MEAN NUMBER OF "I" PLUS "WE" RESPONSES FOR THE SCHIZOPHRENIC GROUP OVER THE FOUR BLOCKS OF 20 TRIALS EACH

critical difference among the means necessary for significance at the .05 level of confidence was calculated. Results of this test showed that Group *RP* performed significantly better than Groups *P* and *C*, and Group *R* attained a significantly higher level of learning than Group *C*. However, the difference between Groups *P* and *C* was nonsignificant. Thus, the greater efficiency of Treatments *R* and *RP* over control conditions for the neurotic group are clearly established, while the reinforcing effect of punishment is ambiguous.

Schizophrenic group. Figure 3 presents the results for the schizophrenic *S*s. In agreement with the results for the normal and neurotic groups, the trials by treatments interaction for schizophrenics was significant ($p < .01$). Calculation of the critical difference among the means necessary for statistical significance at the .05 level revealed that Group *RP* performed significantly better than Groups *R* and *C*, and Group *P* attained a significantly higher level of learning than Group *C*. The difference between Groups *R* and *C* was not significant, however. Thus, the greater efficiency of Treatments *P* and *RP* over control conditions for the schizophrenic group are clearly established, while the reinforcing effect of reward is ambiguous.

Treatment groups. Separate Type I analyses were also applied to each of the four treatment groups in an effort to ascertain whether the frequency of criterion responses differed among diagnoses over blocks. Groups *R*, *P*, and *RP* all showed significant trials effects ($p < .001$), while the control group displayed no such increasing usage of the first-person pronouns during training. Only Group *R* showed a significant trials by diagnoses interaction ($p < .05$), with the schizophrenic *S*s showing a minimal rise in their usage of I and We during the experiment, while normals and neurotics displayed substantial increments. The differences among the Block 4 means of Group *R* were also analyzed by the critical difference technique. Only the difference between the normals and schizophrenics was statistically significant ($p < .05$); the difference between the neurotic *S*s and the schizophrenics did not quite attain significance ($p < .10$). To determine if schizophrenics used I and We significantly more frequently on the fourth block than they did on the first, a t test for related samples was carried out on the data for these two blocks of trials. A $t = 0.65$ ($df = 14$) was obtained, yielding no evidence of an increase in preference for the first-person pronouns during training.

Subjects' hypotheses. As pointed out previously, 14 *S*s were eliminated from the study because they clearly verbalized that their use of one or both of the first-person pronouns de-

termined *E*'s statements. Although *E* had suggested prior to training that the purpose of the experiment was to aid in improving hospital conditions, apparently only three *S*s took this seriously. Almost one half of the 135 *S*s (responses of the control *S*s not included) declared that they were unable to determine what the experiment was about despite close questioning. The next most prevalent notions were that a high speed of reaction was desired by *E* ($N = 20$) and that personality evaluation was *E*'s intention ($N = 20$). Twelve *S*s believed that *E* wished to determine how "self-centered" their thoughts were. The remainder of the experimental group ($N = 15$) looked upon the procedures as a kind of intelligence test.

DISCUSSION

Previous demonstrations of the inducement of learning without awareness are confirmed by the results of the present study. In all three diagnostic groups, significant learning over trials by at least two of the three treatment subgroups occurred, whereas in no case did the control *S*s show this result. Careful questioning of the *S*s failed to indicate any realization that an increment in their usage of the criterion response had occurred in response to verbalizations of approval or disapproval by *E*. Indeed, even the objection raised by Adams (1957) that *S*s' use of correlated hypotheses often explains the experimental results suggestive of learning without awareness does not appear to be crucial in this study. It might be suspected that those *S*s who declared they chose pronouns on the hypothesis that *E* wished to learn of their "egocentricity" would tend to increase in their selection of I and We over trials even though they obviously were aware of this tendency. This would appear to be an example of a correlated hypothesis in Adams's sense since, in fact, 8 of the 12 *S*s who made this hypothesis showed such behavior although not to an outstanding degree. However, since none of the other hypotheses offered would appear to lead to such "correlated" behavior, and since only 12 *S*s shared the above hypothesis, the present results do not seem vulnerable on this basis.

Although learning took place in all diagnoses studied, the relative efficiency of reward and punishment in inducing this learning varied sharply according to diagnosis, often in disagreement with the results of previous investigators. However, it would be well to point out that since no formal hypotheses were tested in the present study, a similar question as to the stability of the present findings can legitimately be raised.

While the normal *S*s showed highly significant learning during the experiment, *E*'s use of reward, punishment, or both, proved equally effective. This equality of performance under the various treatments was not shown by the neurotic and schizophrenic *S*s. Whereas comparisons of each of the treatments with the control group resulted in significant differences in the normal group, the neurotics clearly showed learning only under the combined use of reward and punishment and reward alone. Thus, while the use of reward in the establishment of this response tendency for the neurotic *S*s in this study seems effective, the efficiency of punishment is ambiguous. Cohen and Cohen (1956), in an experiment which utilized reward alone and a task almost identical to the present one, also found that neurotic *S*s displayed significant learning under reward.[2]

The schizophrenic group performed in still a different manner. Although learning also occurred under the combined reward and punishment condition, the psychotic *S*s showed significant learning under punishment but not under reward—just the opposite of the findings with the neurotic group. This failure to learn under reward alone is once more in close agreement with the results of Cohen and Cohen (1956), who found that while neurotics learned under conditions suggestive of approval, schizophrenics did not. The results of Buss and Buss (1956) with a task demanding conceptual learning may be in agreement with these findings. These *E*s also found that reward alone was less effective in inducing learning than punishment alone or punishment combined with reward. However, the degree of consistency between their results and the present findings is open to some question,

[2] This study also varied the reinforced response class to determine whether there was any difference in the learning when third person pronouns were reinforced rather than first person. No difference was obtained.

since the *S*s selected for the above study were described simply as "neuropsychiatric," the authors furnishing no further elaboration as to the patients' diagnoses.

A consideration in the evaluation of the present results which may limit both the generality of the findings and the validity of any comparisons with previous studies is the type of *S* used. The *S*s selected for this experiment had experienced only brief hospitalization, implying that they may have been acutely disturbed for only a relatively short period of time. Therefore, there may very well be systematic differences between such schizophrenic *S*s and those used in previous investigations.

Examination of the treatment groups reveals little differentiation among the diagnostic subgroups in their response to identical treatment. If, in fact, normals, neurotics, and schizophrenics differ in their responsiveness to punishment, the present response measure is too insensitive to indicate any such dissimilarity. Only under the condition of reward does any significant disparity in performance appear, where the schizophrenics learn significantly less (actually showing no significant increase over trials) than the normal *S*s.

SUMMARY

The present study investigated learning without awareness in three diagnostic groups (normals, neurotic and personality disorders, and schizophrenics) when various verbal reinforcing agents were employed to suggest approval, disapproval, or both. *S*s were presented with 80 3 x 5 cards, each of which contained the same four pronouns (I, We, He, They) but a different verb in the past tense. They were instructed to make up a sentence containing the verb indicated on the card and beginning with one of the pronouns. Results were evaluated by analysis of variance techniques according to how the *S*s varied in their use of the first person as a result of reinforcement during the 80 trials. All three diagnostic groups displayed significant learning over trials when *both* reward and punishment were used. Normal *S*s also showed significant learning to reward and punishment when each was given alone. However, neurotics did not learn under punishment alone, whereas schizophrenics did not learn under reward alone. Explanation of the results in terms of *S*s' use of correlated hypotheses was not possible.

REFERENCES

ADAMS, J. K. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, **54**, 383–406.

BUSS, A. H., & BUSS, EDITH. The effect of verbal reinforcement combinations on conceptual learning. *J. exp. Psychol.*, 1956, **52**, 283–288.

COHEN, E., & COHEN, B. D. Verbal reinforcement in schizophrenia. Paper read at Midwestern Psychological Association, St. Louis, 1956.

GREENSPOON, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409–416.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Chicago: Houghton Mifflin, 1953.

Received May 19, 1958.

SUBLIMINAL EFFECTS OF VERBAL STIMULI[1]

GUDMUND J. W. SMITH, DONALD P. SPENCE, AND GEORGE S. KLEIN

Research Center for Mental Health, New York University

EMPLOYING a technique of exposing subliminal forms followed immediately by supraliminal forms, two early experiments have suggested that subliminal forms measurably affect conscious perception. Smith and Henriksson (1955) have reported that when a fan-shaped array of lines was shown subliminally and followed by a clearly perceptible square, the square becomes more trapezoidal, a change consistent with the usual illusion that appears when a square is seen against a fan-shaped background. Klein, Spence, Holt, and Gourevitch (1958) found that different sexual and symbolic figures exposed subliminally provoked different impressions of consciously perceived pictures of people. For example, drawings of ambiguous human figures seemed more masculine when they followed a subliminal male symbol than when they followed a female symbol in another session. However, it was not certain that the meaning of the subliminal figures was solely responsible for these effects since the symbols differed markedly also in shape. Another experiment seemed necessary to test the decisive importance of meaning alone as a subliminal influence upon conscious thought, using stimuli that differed minimally in contour and maximally in meaning.

Words possessing a core of self-consistent, commonly accepted social meanings, but unequivocally different from each other (e.g., ANGRY and HAPPY) meet this requirement. Through the use of words similar in size and printed in capitals, it is possible to present widely different meanings with relatively minimal variation in contour, and to test for the influence of meaning per se upon associated figures. The central hypothesis, then, was that words, exposed below the threshold of recognition, and followed immediately by a clearly supraliminal figure, would influence impressions of the latter. It was further expected, in light of earlier findings (Klein et al., 1958), that *S*s would differ in their sensitivity to the subliminal stimulus. To explore personality correlates of such differences, we also administered the TAT in the present study.

METHOD

Apparatus

An electronically controlled mirror (Dodge) tachistoscope was used which consists of three fields: a blank field on which *S* is ordinarily focused and the two stimulus fields.[2] Each stimulus compartment is equipped with two mercury argon tubes surfaced with a special phosphor coating that produces a daylight white color, with rapid onset and decay so that the light output closely approximates a square wave form even at very brief exposures.

The illumination of each field can be independently varied. In the present study, all fields were adjusted to equivalent brightness by means of a photometer placed at the viewing aperature. The brightness value of the three fields was fixed at .8 apparent foot candles for the entire range of exposure times used (as determined by MacBeth Illuminometer). The fields were arranged in the box at a 90° angle from each other, at a distance of 3.5′ from the eyepiece. Coated low-reflection mirrors with a transmission-reflection ratio of approximately 50–50 are fixed in the box so as to transmit the two stimuli over a common optical path to *S*'s position. The placement is such as to equate the retinal sizes of each of the stimulus fields. The two stimulus fields may be exposed in succession with no intervening interval; each can be varied from .002 to 5.9 seconds.

Method of Presenting Subliminal Stimuli

A line drawing of a neutral and expressionless face (Fig. 1) was placed in the blank field of the tachistoscope and was clearly visible during all but brief intervals. The words ANGRY and HAPPY, each occupying one of the other two fields, were exposed individually during these intervals in a partially alternating series of increasing exposures. As *S* looked at the face, a word

[1] This study was carried out under a grant from the National Institute of Mental Health. Patients of Manhattan VA Hospital served as *S*s and we are grateful to Bernard Locke, Chief Psychologist, Manhattan VA Hospital, for generously helping to make available the hospital's facilities.

H. Smith, Helen Loomis, Reeva Safrin, Harriette Weintraub, Lester Alston, Sheldon Bach, Morris Eagle, and Leo Goldberger, all of the Research Center for Mental Health, assisted as judges at different points in the experiment. We are especially grateful to Robert R. Holt for valuable counsel in the analysis of TAT protocols and for a critical reading of this manuscript.

[2] A complete description of the tachistoscope and electronic timing units may be obtained from Charles Snyder, The Menninger Foundation, Topeka, Kansas, who designed and constructed them.

FIG. 1. SUPRALIMINAL STIMULUS

was briefly exposed (less than 40 msec.), and then the face immediately reappeared. In the next cycle either the same or the other word again interrupted the appearance of the face. An electronic switching apparatus controlled the exposure and alternations of each word independently without making *S* aware of the changes.

The subliminal stimuli (ANGRY and HAPPY printed in block capitals) were shown at gradually lengthened exposures. Although the face disappeared momentarily whenever a word came on, none of the *S*s consciously detected the word itself. The face seemed only to flicker momentarily. We assumed that if the word had indeed registered, its meaning would influence reactions to the associated face.

Procedure

Experimental sessions each lasted about one hour and took place on successive days. Since tachistoscopic acuity varies enormously, it was necessary to determine in a preliminary session each *S*'s sensitivity to words under conditions similar to the main task. Such a session seemed also advisable to put *S* at ease in this unfamiliar situation.

Session I: The first session began with the following instructions:

> We're making up a series of visual tests and we'd like your help in putting them together. We're interested in different kinds of visual sensitivity and we hope to see how useful some tasks may be for our purposes. We'd like your reactions to everything we do.

S then gave stories to three TAT pictures: Cards 1 (boy with violin), 8BM ("operation" scene), and 12BM ("hypnosis" scene). Rapaport's (1945) instructions for the TAT introduced each card, and *S*s' responses were recorded on a Gray Audograph.

S then moved to the tachistoscope. He sat in a canvas chair, back and head pillowed, and gazed into the box. A black curtain, enveloping his head, reduced outside visual distractions. He saw before him a landscape—a colored photograph of a Belgian town square from the *National Geographic Magazine*. Off to one side of the town square was a dominating tall monument, with several people seated at its base. Ringing the square was a group of stone houses. Blue sky filled the upper third of the picture, and it was in this area that the subliminal word appeared. *S* was told:

> Here's the first task we want to get information about. Make yourself comfortable and look into the box. Each time I say "Ready," tell me something about what you see—something you hadn't noticed before, or some part which appeals to you. Don't worry about being too accurate—just tell me something each time I say, "Ready." We want to see how observant you are. The new part may be in the picture, come after the picture, or even come before the picture.

To determine *S*'s threshold for the "masked" word, one of five words (HABIT, MOTOR, BIRTH, ADOPT, SCALE) was momentarily exposed, interrupting the appearance of the landscape. All five words were equated in length and frequency to ANGRY and HAPPY on the Thorndike-Lorge word count (Thorndike & Lorge, 1944). They were randomly assigned to the *S*s. *S*'s recognition threshold was the exposure time at which he saw letters or simply mentioned that something like a word had appeared. An occasional *S* reported that he had actually seen the word earlier than his first mention of it. In such instances the recognition level was adjusted downward. The median recognition threshold for the sample was 125 milliseconds.

S's threshold for the "masked" word in the first session determined the initial subliminal exposure in the main experimental condition. If the recognition threshold for letters or word was 28 to 80 milliseconds, the initial exposure of HAPPY was 4 milliseconds. If the initial threshold was higher than 80 milliseconds, the initial exposure for HAPPY was raised accordingly (see note on Table 1). The schedule of exposures was planned to allow for presenting ANGRY and HAPPY 21 times each at gradually increasing exposures before reaching *S*'s recognition region. Not all *S*s were shown the two words the same number of times because four *S*s recognized one of the words before the final exposure. However, each stimulus word was given at least 13 times to all *S*s, and 17 times or more to all but one *S*. Initial exposure of the subliminal HAPPY varied from 4 to 11 milliseconds, and final exposures of the subliminal ANGRY varied from 20 to 70 milliseconds for all but one *S* whose range was 45 to 130 milliseconds. The final exposure was usually below the threshold level determined in the first session. We eliminated from analysis all responses to the face made after the subliminal word was recognized.

The sequence and the most frequently used range of exposures are given in Table 1. The relative increase in exposure time from one exposure to the next was the same for all *S*s, a difference between the logarithms of the two exposure times of approximately .07. Stimuli appeared in the same order for all *S*s (see Table 1).

Session II: S sat at the tachistoscope in the manner described above. The instructions were:

> This time we want you to look into the box and describe the face you see inside. Tell us everything that comes to mind about this person, just the way

TABLE 1
SEQUENCE AND EXPOSURE OF SUBLIMINAL STIMULI

Exposure Number	Subliminal Word	Modal Exposure (milliseconds)[a]
1, 2, 3, 4	HAPPY	4
5, 6, 7, 8	ANGRY	5
[9—image]		
10[b]	ANGRY	5
11, 12, 13, 14	HAPPY	6.5
[15—image]		
16, 17, 18, 19	ANGRY	7
20, 21, 22, 23	HAPPY	8
[24—image]		
25[b]	HAPPY	8
26, 27, 28, 29	ANGRY	10
[30—image]		
31, 32, 33, 34, 35[b]	HAPPY	11
36, 37, 38, 39	ANGRY	14
40, 41, 42, 43	HAPPY	16
44, 45, 46, 47	ANGRY	20

[a] Nine *Ss* were shown the stimuli at exposures indicated; 4 *Ss* at levels 2 milliseconds higher; 5 *Ss* at levels 4 milliseconds higher; and 1 *S* at levels 7 milliseconds higher. Increases were determined by *S*'s threshold in first session.

[b] A fifth exposure was added occasionally to interrupt the regular sequence, but responses to these exposures were not tabulated in order to preserve the symmetrical design.

you described the pictures we showed you yesterday. Try to make up a story about him. Tell us what he does, how old he is, what he is thinking about, and who he reminds you of.

After describing the face for two minutes, *S* was instructed as follows:

The next task has to do with how sensitive and accurate people are to changes in other people's faces. As you know, this is a very important ability. We are going to show you the face you just saw. As you watch we will change the expression in the face. The changes we make will usually be small and subtle, and we want you to tell us in what ways the expression changes. At first, the changes may be hard to pick up, but gradually you will see more and more change. Tell as much as you can each time. When I say "Watch," the face will change. Tell us how it changed—in what ways it looks different. Sometimes you may sense a change in mood, even though the face may look the same. Tell us about that, too. From time to time we will ask you to move back, close your eyes, and form an image the way you did just now [in a preceding practice session]. We'll tell you when the time comes.

All responses were recorded on Audograph discs. The test procedure was carried out by two *Es*; one controlled the stimulus presentations and exposure times; the other sat behind and to the side of *S*, away from the control panel, and was therefore ignorant of the actual word being presented. His job was to question vague or ambiguous responses and to prompt *S* after long silences or to question him if *S* had reported "no difference" over a period of five or more exposures. In this arrangement *E*'s questions to *S* were free of possible bias from knowledge of the subliminal stimulus operating at the moment. The partial aperiodic sequence of stimuli (see Table 1) gave further assurance that in conducting his inquiry *E* could not easily guess the identity of the stimulus.

In addition to the recorded verbalized responses, the *time* taken for *S* to respond was also recorded. Latency was measured by stop watch from the time the stimulus appeared until *S* began to speak.

At four irregular intervals (Exposures 9, 15, 24, and 30) *S* was asked to lean back, close his eyes, and form an image. His descriptions were recorded. Two of these intervals followed a block of HAPPY exposures and one followed a block of ANGRY exposures (see Table 1). We had hoped that *S*'s image of the face would also reflect the subliminal processes activated by the stimulus words, but most *Ss* seemed unable to form images easily and for this reason, these data are omitted from the analysis.

At the close of the experiment *S* was instructed not to discuss the experiment with the other patients who served as *Ss*. There is no guarantee that *Ss* obeyed this instruction, but there was no evidence that any of them disobeyed, and since so few actually recognized the subliminal word it is unlikely that much vital information made the rounds of the wards.

Subjects

The 20 *Ss* were male patients from open wards of a VA hospital. They ranged in age from 21 to 55, with a median age of 40. *Ss* were selected only if they seemed moderately fluent and capable of expressing themselves verbally, and if they were not under the influence of tranquilizers or depressant drugs. The symptom picture was varied, with 6 *Ss* complaining of ulcers or disturbances of the gastrointestinal tract, 4 of depression, and 4 of anxiety and mixed conversion symptoms. There were 2 asthmatics, 2 mixed paranoids, and 2 not classified. Occupations were generally semiskilled, with the exception of one magazine editor, one recording engineer, and one jeweler. Testing was made part of each patient's hospital routine, but every effort was made to have the patient understand it had no relevance to his stay at the hospital.

RESULTS

Response Classifications

Before scoring responses given in the critical experimental condition, we divided *Ss* randomly into two groups of ten, one of which was used to develop categories for classifying the widely varying descriptions of the face. One set of categories referred to *expressive qualities* of the face, e.g., mood, character,

thoughts. Another set classified *physical features* such as asymmetry, shading, shape of eyes and nose, roundness of face, etc. Since only the positive and negative expressive categories bear directly on the hypothesis, we will confine our presentation to these categories. They were as follows:

Pleasant: Pleasant frame of mind; positive affect: e.g., happy; smiling; laughing; cheerful; pleasant; relaxed; liking something. This category was used whenever there was evidence of "smiling," "contentedness," etc., so long as it was uncontaminated by negative feelings.

Complex-Pleasant: Complex frame of mind with pleasant overtones or predominantly positive characterization: e.g., happy with a serious shadow; false and phony smiles; sneery smile; just a pleasant front; withholding a smile; and negations of positive characterization: e.g., not happy.

Unpleasant: Rejecting, hostile frame of mind; negative affect and mood states: e.g., angry; mad; anxious; depressed; sad; moody; sneering; frowning; stern; unpleasant; cruel; sadistic; disliking; hard; cold.

Complex-Unpleasant: Complex frame of mind with unpleasant overtones: e.g., mad but also a little bit smiling; cynical; hostile with a smiling left eye; and negations of negative terms: e.g., not unpleasant; not angry.[3]

*S*s who were reluctant to describe the face or who seemed to break off a description abruptly were prompted now and then. In a later count, these instances of inquiry were found to be randomly distributed in the group over the entire series. Occasionally *S*'s responses to a question produced new information or differed somewhat from his spontaneous description. In the content analysis, all of the material produced by *S* after any one exposure was considered together, whether or not part of it was elicited by such questioning.

The following excerpts from the protocols of two *S*s illustrate how the scoring categories were used. The number before each response refers to its position in the series. The subliminal word and the scoring category applied to the response are in brackets. The response "same as previous picture" was given the score of the preceding response.

Subject A:

1 (HAPPY): There's a slight change, a little grin, like a smile, starting to smile (Pleasant).
2 (HAPPY): Looks the same as the first picture (Pleasant).
3 (HAPPY): Seems, uh, the left eye seems more almond-shape than the right eye. The right eye seems more round (not scored for expression).
4 (HAPPY): The left eye seems more almond-shape than the right. The right eye seems more round (not scored for expression).
5 (ANGRY): It seems that the left eye isn't almond-shape as before. The right eye is still round (not scored for expression).
6 (ANGRY): I can't see much difference. I saw the eye. That's about all (No Difference).
7 (ANGRY): Looks like the first picture and looks like he's thinking of, far away, just uh, daydreaming (Pensive—not scored for affect).
8 (ANGRY): No, I don't see much difference in this picture than in the last one (Pensive).

Subject B:

1 (HAPPY): Well, eh, the description of the face seems to be the same, and, eh, the shadow on his, the right side, seems to have gotten lighter (No Difference).
2 (HAPPY): I didn't notice much change that time (No Difference).
3 (HAPPY): I still haven't noticed any change (No Difference).
4 (HAPPY): He seems to be smiling a little bit now (Pleasant).
5 (ANGRY): His face seems to be brightened up now. Away from that frowned expression (Complex-Pleasant).
6 (ANGRY): No change. (Q) The outline—the heavier outline on his nose now. More so on his left side and partially on his right (No Difference).
7 (ANGRY): He is getting that moody look again (Unpleasant).
8 (ANGRY): The shadow on the left side of his face seems heavier (not scored for expression).

[3] Other categories used in scoring, but not analyzed in this paper, were as follows:

Pensive: A serious-thoughtful frame of mind: e.g., daydreaming; waiting; thoughtful; not concentrating; considering; sleepy; contemplative; serious (within the connotation of thoughtfulness).

Active: An outward-directed frame of mind: e.g., alert; observing; staring; looking steady; decided, ready to do something.

Complex: Complex frame of mind, neither entirely Pleasant nor Unpleasant: e.g., not angry, not happy; one eye mad, one eye cheery.

Puzzled: e.g., questioning; surprised; undecided; uncertain; problem on his mind; bewildered; concerned; distracted; not knowing; quizzical; worried.

Neutral: Inexpressive or blank: e.g., just normal; nothing special; just looking; empty; vacant; nothing in him.

No Difference: This was scored when *S* saw no change in the face, except when it was called "the same as the previous picture." Then it was scored in the category used for that picture.

TABLE 2
TOTAL POSITIVE (+) AND NEGATIVE (−) REACTIONS TO THE SUBLIMINAL STIMULI

	HAPPY			ANGRY			
S	+	−	d_1*	+	−	d_2	$\frac{d_1}{d_2}$*
1	2	0	2	5	2	3	−1
2	0	1	−1	0	1	−1	0
3	4	0	4	0	0	0	4
4	4	0	4	2	0	2	2
5	1	1	0	1	1	0	0
6	6	2	4	1	2	−1	5
7	5	2	3	3	2	1	2
8	4	2	2	3	1	2	0
9	2	0	2	2	0	2	0
10	0	0	0	0	0	0	0
11	0	1	−1	0	3	−3	2
12	1	0	1	1	3	−2	3
13	2	0	2	1	0	1	1
14	5	0	5	1	0	1	4
15	0	0	0	0	0	0	0
16	1	0	1	0	0	0	1
17	4	0	4	4	0	4	0
18	11	0	11	8	1	7	4
19	3	0	3	3	2	1	2
20	9	0	9	3	0	3	6

* $p < .005$

Three judges classified the responses of the first 10 *S*s in the above categories. The few instances of disagreement among them were easily resolved by tightening and clarifying definitions of the categories. Then all 20 cases were scored on the revised categories. Because of the high agreement among judges on the first sample, only two judges were used for the second scoring. Whenever possible, judges assigned a response to only one of the expressive categories. If that proved difficult, the response could be placed in several categories, but only the predominant category was scored. Four responses in each series were scored, the occasional fifth exposure being ignored.

Subliminal Effects on Verbal Descriptions

Verbal stimuli were considered subliminal if they fell below the threshold of *spontaneous* report in either the first or second session. In neither case was *S* asked to look for a word or letters. When four of the 20 *S*s reported "letters" toward the end of the experimental series we discarded all responses at that exposure setting and, if necessary, balanced the trials by excluding responses to the preceding, alternate stimulus. Our final data, therefore, consist of reactions to a supraliminal face accompanied by verbal stimuli outside the range of *S*'s spontaneous report. We are asking, then, whether or not words have an effect when *S* does not report their presence.

Responses to HAPPY and ANGRY were tallied separately (see Table 2). In the HAPPY condition, 15 *S*s described the face in a manner more Pleasant or Complex-Pleasant than Unpleasant or Complex-Unpleasant; 2 *S*s found the face relatively more Unpleasant; and 3 *S*s either did not sense any expression or found the face equally Pleasant and Unpleasant. The stimulus HAPPY, therefore, produced significantly more Pleasant than Unpleasant reactions to the face ($p < .005$, Wilcoxon Signed Rank Test, one-tail).

In the ANGRY condition, responses were not significantly more Unpleasant; only 4 *S*s responded with more negative than positive responses as compared with 11 *S*s with more positive responses. If, however, we compare the two conditions (see Table 2, last column) we find that the predominant reaction to the stimulus HAPPY was *more* positive than to the stimulus ANGRY ($p < .005$, Wilcoxon Signed Rank Test, one-tail), a result that supports our main hypothesis.

Impressions of a face, then, are influenced by stimuli outside the range of *S*'s awareness. Some *S*s *may* have been partially aware of the stimulus before they reported it. To what extent do our results depend on partial awareness of the two words? As a control, we presented the face and the two stimulus words to five additional male *S*s (students in an undergraduate psychology class) with the experimental instructions to watch for changes in the face and these additional words:

> Sometimes, in addition to the face, you may see a word—there are definitely words there. But at first they will be flashed too quickly for you to see them. However, be prepared.

After each exposure, *S* was asked, "What was the expression?" and "What word do you think was there?" The stimulus words were presented in a gradually increasing alternating series beginning at 4 milliseconds, the lowest exposure used in the main experiment, and continuing until both words had been correctly recognized on two successive exposures. Illumination was fixed at the value used in the main experiment.

No *S* recognized either letters or a partial word at 4, 5, 6, 7, or 8 milliseconds. At 10 milliseconds, one *S* saw ANGRY and three of the five *S*s saw the outline of a word without being able to identify it. This threshold is substantially lower than the lowest pretest threshold in our sample (28 milliseconds), and reflects the sensitizing effect of instructions to look for words. Returning to the experimental data, we selected those *S*s who had been exposed to both stimulus words at least 16 times below the 10 millisecond level. We compared their responses below 10 milliseconds with their first 16 responses above 10 milliseconds.

Both above and below the 10 millisecond point, reactions to the face were relatively more pleasant when the word HAPPY was exposed than when ANGRY was exposed, with the effect slightly greater *below* 10 milliseconds. Moreover, the first 16 responses of these nine *S*s to words below 10 milliseconds did not differ significantly, in terms of appropriate pleasantness, from the first 16 responses of the remaining *S*s. It would seem, then, that the effects of HAPPY and ANGRY do not depend on exposure level, because responses to the face below a 10 millisecond exposure of the first figure are not significantly different from responses above 10 milliseconds. The low recognition threshold in the control trials suggests that awareness of the stimulus is possible at low levels given an instruction that alerts *S* to the presence of the stimulus; hence, the possibility is not ruled out that some of the experimental *S*s may have been at least partly aware of the stimulus before reporting it. However, the absence of difference between responses at low and high exposures suggests that information gained from partial awareness of the stimulus cannot be a necessary condition for the effect. Because the influence of the words does not increase as exposure increases, it appears that the sheer amount of consciously available information provided by the A-exposure is *less* important than other variables, as yet unspecified.

A sequential analysis supports the finding that reactions to HAPPY and ANGRY are not a function of duration of the "subliminal" word alone. Responses were predominantly Pleasant or Complex-Pleasant in the first and last blocks of HAPPY exposures (4 exposures in each block), but predominantly Active, Neutral, Puzzled, and Complex in the middle block. Unpleasant responses were more frequent in the *first* ANGRY block than in either the middle or final blocks. Nor does the effect of each word increase within each block of exposures. In 8 out of 10 blocks of HAPPY exposures (with the exception of the last 2), Pleasant and Complex-Pleasant responses usually came at the very first exposure. On the other hand, Unpleasant and Complex-Unpleasant usually came on the last exposure of ANGRY within a block.

There is evidence that *S*'s first impressions of the face—positive or negative—given *before* the tachistoscopic trial tended to affect the influence exerted by the subliminal words. Nine *S*s initially described the face as predominantly Pleasant. Among the other 10 *S*s (the 11th was excluded because he saw no change in the face) one was uncommitted, and the remaining 9 found the face at least partly Unpleasant. Only two *S*s of the first group (Pleasant initial impression) used the Unpleasant category during the entire experiment; 8 *S*s of the last group used it frequently. The difference between these proportions is significant ($p < .02$, Fisher's Exact Test for 2×2 contingency tables). Evidently, initial impressions set the theme for subsequent reactions to the face, yet preference for one type of response did not rule out distinctive reactions under the two conditions. Thus, one *S* who saw the face as initially Pleasant found it Complex-Pleasant when paired with ANGRY.

Subliminal Effects Upon Report Latency

In order to normalize the distribution of latencies among *S*s and exposures, time values were converted to logarithmic units. When all latencies were averaged, there was no perceptible trend favoring either the HAPPY or the ANGRY series. About half the *S*s (9 out of 20) had longer latencies for the HAPPY series. However, when the series were analyzed separately, it appeared that latency in the ANGRY pairings was longer when *S*s gave Unpleasant and Complex-Unpleasant responses than when they gave Pleasant and Complex-Pleasant responses. Eight *S*s used both groups of categories in their responses to ANGRY, and for six of them, this trend was evident ($p < .05$, Wil-

coxon Signed Rank Test, two-tail). In the HAPPY series, there was no tendency to delay negative reactions more than positive ones.

The hesitation before giving negative responses to the ANGRY face fits well with the tendency to give negative responses late within each block of ANGRY pairings. Our *S*s gave negative responses only with perceptible reluctance, which probably reflects their unwillingness to describe the face in anything but positive terms, as indicated by a preponderance of positive responses in both series. The ANGRY stimulus, if acknowledged, required the use of a new set of responses which took time to mobilize.

With the positive descriptions, long latency did not accompany a particular type of response; on the contrary, the impact of the subliminal HAPPY tended to appear *either* in latency reaction *or* in a descriptive response, but not in both. The rank order correlation between latency difference and preference for positive responses was $-.54$ ($p < .05$). This finding means that *S*s who used more Pleasant and Complex-Pleasant responses in the HAPPY than in the ANGRY pairings did not differ in their reaction times to the two series, whereas *S*s who differed in reaction times (having either a longer or a shorter latency for the HAPPY face) used the positive response categories equally often in both conditions.

TAT Indications of Subliminal Sensitivity

Despite the strong general trends, it was expected that the impact of the subliminal stimuli would vary among *S*s. To assess these variations, we examined the TAT records obtained in the first session. Several methods of analysis were attempted; here we will report only the final investigation which was based on an earlier sampling of the stories for significant signs. Signs that proved differentiating in the scoring of two judges were grouped into three clusters and expressed as three bipolar rating scales. To these we added three scales that had discriminated Sensitive and Insensitive *S*s in a parallel study (Bach & Klein, 1957). All six scales[4] were then given to a new pair of judges not acquainted with previous results who were asked to apply them to the TAT stories.

The judges independently distributed the *S*s by dividing them into four groups of five *S*s each on each scale. Differences in each four-fold division were reconciled by discussion between the judges. The three stories for each *S* were considered together in arriving at his position on any one scale. Each *S*'s TAT score was then compared with his sensitivity to the subliminal stimuli, that is, to his ability to give a response congruent with HAPPY or ANGRY. Subliminal sensitivity was determined by means of a matching model (Mosteller & Bush, 1954, p. 310) which provides a significance level for each *S*. Scores ranged from highly Sensitive (better than chance) to Insensitive (poorer than chance).

Two scales were found to be related to subliminal sensitivity: *Story departs from immediate details of picture* vs. *Story follows closely on picture*, and *Main characters passive in posture or attitude* vs. *Main characters dominant and rational, not given to daydreaming or fantasy.* On the first scale, Sensitive *S*s tended to tell a TAT story which adhered to the details of the picture ($.05 < p < .10$, two-tailed, Wilcoxon Test for Differences in Location). These *S*s told stories in which motivation and character were often suggested by expressions of posture of the pictured figures, and any deviations added only with caution. For example, one Sensitive *S* described Card 1 in this fashion: "Looks like he's studying something—it's a violin. Just looking at the way it's made. Looking at the bows. Do you call it bows? Looks like he's looking down." Insensitive *S*s tended to give more elaborate stories and

[4] The six scales were as follows: (*a*) *S* not certain of story or outcome, gropes for words vs. *S* shows little doubt about pattern or outcome of story; fluent. (*b*) Stories loosely structured, may contain contradictions; *S* aware of many possibilities vs. Stories adequately structured with no ambivalence or contradiction, single story line. (*c*) *S* identifies positively with one or more characters vs. *S* shows little sympathy with any character, and characters unconcerned about one another. (*d*) Story departs from immediate details of picture; new elements introduced; picture details elaborated vs. Story follows closely on picture with no new elements introduced. (*e*) Main characters passive in posture or attitude vs. Main characters dominant and rational, not given to daydreaming or fantasy. (*f*) Story outcome positive or hopeful vs. Story outcome negative or pessimistic.

supply motivation and detail, as in this example: "Well, that's a youngster — seems to be a tomboy over a violin. From the expression on his face, he's not so sure he wants to play the violin, he'd probably rather have something a little noisier, possibly a clarinet, cornet, or even the drums." Another Insensitive *S* on the same card found the boy "living in a dream of what might be. He'll want to be a railroad engineer, a fireman, a magician."

It is significant that although Sensitive *S*s were faithful to the pictured details, they were not restricted to them, whereas *S*s who adhered more closely to the picture were only moderately sensitive to HAPPY and ANGRY. Apparently Sensitive *S*s showed the effect because they were both attentive and flexible in their responses. Insensitive *S*s, on the other hand, either adhered rigidly to the supraliminal face and so gave little room for the influence of HAPPY or ANGRY, or else they were preoccupied by their own fantasies and described the face in a way that had no relation to the subliminal stimuli.

Acceptance of passivity was the other variable found to distinguish Sensitive from Insensitive *S*s. In the original ranking on *Main characters passive*, stories by both extreme groups were found to stress passivity in their characters, but further inspection showed that Sensitive *S*s were also more *tolerant* of passive attitudes and more optimistic about the consequences of submission. For example, the patient in Card 8 was described by one Sensitive *S* as a man who "is thinking of getting operated on. Looks like his future is pretty good because he's all dressed up and he's real clean. Looks like he's doing all right." Insensitive *S*s, although they emphasized passivity, were more suspicious of its outcome and distrustful of the authority figure who was enforcing the passivity. Thus one Insensitive *S* was reminded by Card 12M of Svengali and Trilby; another Insensitive *S* was quick to deny threat on Card 12M and said, "It's almost a threatening gesture but not quite. . . . Yes, a little bit too relaxed to be threatening." When it became apparent that acceptance of passivity was the critical variable, the five extreme Sensitive *S*s and the five extreme Insensitive *S*s were ranked on this variable by two new judges as a check. Four out of five Sensitive *S*s were judged more tolerant of passivity as compared to only one out of five Insensitive *S*s ($p = .10$, Fisher Exact Test).

In general, one might say that the passive role forced on them in the experiment had special significance for both sets of *S*s, and that their response to the subliminal stimulus depended on whether they tolerated or were suspicious of this role. In this particular experimental situation, tolerance of passivity appeared to accompany a flexibility of report, whereas suspiciousness accompanied either a rigid, unyielding response or preoccupation with personal fantasies about the face.

DISCUSSION

The present study, like others which preceded it (Klein et al., 1958; Smith & Henriksson, 1955), has shown that registration outside awareness can produce alteration in conscious impressions of an associated stimulus. The unique feature of the present study was the use of words as subliminal stimuli which were more obviously different in meaning than in contour. The words HAPPY and ANGRY shown subliminally resulted in different impressions of a consciously seen face. It seems clear, therefore, that meanings were aroused by the subliminal words, and that these meanings directly affected responses to the contiguous conscious stimulus.

While the changes in the face are impressive and in the predicted direction, the cause of these changes is not yet clear. The fact that the A-stimulus was not reported during the experimentl trials suggests that it influenced the perception of the face while being outside awareness. It seems fair to assume, however, that some *S*s had some partial awareness of the stimulus because when attention was directed to the words (in the control series using five *S*s) words were recognized. Possibly the awareness was of a sporadic nature, and too fragmentary to be readily verbalized during the session. But partial awareness alone does not seem to explain our results; the effect of HAPPY and ANGRY does not increase as exposure time increases. Changes in the face were *not* more pronounced at the more prolonged exposures of the two stimuli. On the contrary, the effect of the subliminal stimuli was more pronounced at the beginning than in the intermediate trials. Put another way, an *S*'s chances of describing the face in accordance with the A-

stimulus were not significantly greater in the later series of exposures than in the earlier series.

It would appear that some other concept is necessary to account for our findings, and we suggest *awareness* as one possibility. *Lack* of awareness, we assume, increases the effectiveness of the A-stimuli, because as peripheral stimuli they escape the disciplined categorizing and appraisal of conscious judgment which fosters differentiation of stimuli and they can more easily be combined with impressions of the face. Presented in awareness, the stimulus might be more easily viewed as a word followed by a face, and impressions of the first might not be assimilated to the second. Hence the stimulus would have less effect on the face in full awareness than out of awareness. This hypothesis is currently being tested in a further experiment in our laboratory, in which we are comparing the effects of concurrent subliminal and supraliminal stimulation.

From the present data, it would seem that *S* is primarily aware of the face during the experimental session because he is instructed to watch for changes in expression. These instructions make him only marginally aware of any peripheral stimulation, and the very short A-exposures further reduce the chance of his recognizing the words. As was shown in the control study, *S* is *capable* of recognizing the words at fairly brief exposures if he is alerted to them, and this occurred, on the average, at an exposure briefer than many of the experimental trials. Hence we might say that during a large part of the experiment proper, *S* was *fully* aware of the face and *potentially* aware of the words. *S* may have been responding to unverbalized fragments of the words, but at the time he was not consciously aware of their presence.

It seems safe to assume that HAPPY and ANGRY may have registered, in whole or in part, without reaching awareness because *S* was not set to discriminate them. We would suggest, tentatively, that *because* they did not reach awareness they were relatively effective in influencing reactions to the face.

The effect of the subliminal stimulus may show itself in both verbal report and response latency, and there is evidence that latency is delayed if *S* is uncomfortable expressing the change verbally. Negative responses to ANGRY were delayed more than positive responses, and came late in each block of four exposures, suggesting that *S*s were reluctant to respond to the face in negative terms. One reason for this reluctance may have been that the majority of *S*s viewed the face as initially pleasant; hence, a response of "unpleasant" was a definite shift in reaction which would take time merely to express. Repeated stimulation by the subliminal ANGRY seems necessary before its meaning can be assimilated to the reaction to the face. Thus it would appear that in addition to being out of awareness, the subliminal stimulus must be compatible in meaning with the supraliminal stimulus. If the two are decidedly disparate, they may be difficult to reconcile verbally and result in uncertainty which is expressed by the report "No difference," or, in less extreme cases, in delayed response.

In general, exposure of a subliminal stimulus furnishes *S* with a set of peripheral meanings that are outside of awareness but available for assimilation to the consciously seen face. If the stimulus is compatible in meaning with the face, and if *S* is able to assimilate the two meanings his reaction to the face should be in accordance with the meaning of the stimulus. Ability to assimilate is apparently related to tolerance for a passive role, and to the ability to maintain moderate distance from the conscious stimulus so that *S* can describe it faithfully without being pedantic or literal.

SUMMARY AND CONCLUSIONS

The central hypothesis tested was that the meaning of words, exposed below the threshold of recognition, and followed immediately by a clearly supraliminal figure, would influence impressions of the associated stimulus.

The words HAPPY and ANGRY were presented tachistoscopically in a mixed sequence of gradually increased subliminal exposures, each alternating with a drawing of a relatively expressionless face. The face was clearly visible for all but the brief interval when HAPPY or ANGRY was flashed on the exposure field.

Twenty male patients (median age 40 years) on open psychiatric wards of a Veterans Administration hospital served as *S*s.

When *S*s' descriptions of the face were classified, they were clearly more Pleasant in HAPPY pairings than in ANGRY pairings. The Pleasant

effect seemed to appear early in HAPPY pairings; the Unpleasant effect late in ANGRY pairings. The initial description of the face tended to modulate the form of influence exerted by the subliminal words. If the effect of the subliminal stimulus was not discernible in the comments, it tended to appear in the time which elapsed before *S* responded to the face. An analysis of TAT records revealed that *S*s who were most and least affected by the subliminal words differed with respect to tolerance for passivity and flexibility of report.

The main conclusion is that differences in meaning between words registered below the threshold of recognition can affect associated conscious thoughts, although these influences are modulated by response preferences and certain idiosyncratic reaction tendencies.

REFERENCES

BACH, S., & KLEIN, G. S. The effects of prolonged subliminal exposures of words. Paper read at American Psychological Association, New York, September, 1957.

KLEIN, G. S., SPENCE, D. P., HOLT, R. R., & GOUREVITCH, SUSANNAH. Cognition without awareness: Subliminal influences upon conscious thought. *J. abnorm. soc. Psychol.*, 1958, **57**, 255–266.

MOSTELLER, F., & BUSH, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. 1. Cambridge, Mass.: Addison-Wesley, 1954.

RAPAPORT, D., GILL, M. M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945.

SMITH, G. J. W., & HENRIKSSON, MAJ. The effect on an established percept of a perceptual process beyond awareness. *Acta Psychol.*, 1955, **11**, 346–355.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Teachers Coll., Columbia Univer., 1944.

Received May 29, 1958.

THE EFFECT OF SEVERITY OF INITIATION ON LIKING FOR A GROUP[1]

ELLIOT ARONSON
Stanford University
AND JUDSON MILLS
U. S. Army Leadership Human Research Unit, HumRRO

IT IS a frequent observation that persons who go through a great deal of trouble or pain to attain something tend to value it more highly than persons who attain the same thing with a minimum of effort. For example, one would expect persons who travel a great distance to see a motion picture to be more impressed with it than those who see the same picture at a neighborhood theater. By the same token, individuals who go through a severe initiation to gain admission to a club or organization should tend to think more highly of that organization than those who do not go through the severe initiation to gain admission.

Two questions are relevant here: 1. Is this "common observation" valid, that is, does it hold true when tested under cotrolled conditions? 2. If the observation is valid, how can it be accounted for? The relationship might be simply a result of differences in initial motivation. To take the case of initiations, persons who initially have a strong desire to join a particular club should be more willing to undergo unpleasantness to gain admission to it than persons who are low in initial interest. Therefore, a club that requires a severe initiation for admission should be joined only by those people with a strong desire to become members. On the other hand, a club that does not require a severe initiation should be joined by some individuals who like it very much, and by others who are relatively uninterested. Because of this self-selection, one would expect persons who are members of clubs with severe initiations to think more highly of their club, on the average, than members of clubs without severe initiations.

But is there something in the initiation itself that might account for this relationship? Is severity of initiation positively related to group preference when motivation for admission is held constant? Such a relationship is strongly implied by Festinger's (1957) theory of cognitive dissonance. The theory of cognitive dissonance predicts this relationship in the following manner. No matter how attractive a group is to a person it is rarely completely positive, i.e., usually there are some aspects of the group that the individual does not like. If he has undergone an unpleasant initiation to gain admission to the group, his cognition that he has gone through an unpleasant experience for the sake of membership is dissonant with his cognition that there are things about the group that he does not like. He can reduce this dissonance in two ways. He can convince himself that the initiation was not very unpleasant, or he can exaggerate the positive characteristics of the group and minimize its negative aspects. With increasing severity of initiation it becomes more and more difficult to believe that the initiation was not very bad. Thus, a person who has gone through a painful initiation to become a member of a group should tend to reduce his dissonance by over estimating the attractiveness of the group. The specific hypothesis tested in the present study is that individuals who undergo an unpleasant initiation to become members of a group increase their liking for the group; that is, they find the group more attractive than do persons who become members without going through a severe initiation.

METHOD

In designing the experiment it was necessary to have people join groups that were similar in every respect except for the severity of the initiation required for admission—and then to measure each individual's evaluation of the group. It was also necessary to randomize the initial motivation of subjects (*S*s) to gain admission to the various groups in order to eliminate

[1] This research was partially supported by a grant from the National Science Foundation, administered by Leon Festinger. The authors are grateful to Leon Festinger for his help and encouragement during the planning and execution of the study.

systematic effects of differences in motivation. These requirements were met in the following manner: Volunteers were obtained to participate in group discussions. They were assigned randomly to one of three experimental conditions: A *Severe* initiation condition, a *Mild* initiation condition, and a *Control* condition. In the Severe condition, *S*s were required to read some embarrassing material before joining the group; in the Mild condition the material they read in order to join the group was not very embarrassing; in the Control condition, *S*s were not required to read any material before becoming group members. Each *S* listened to the same tape recording which was ostensibly an ongoing discussion by the members of the group that he had just joined. *S*s then evaluated the discussion.

The *S*s were 63 college women. Thirty-three of them volunteered to participate in a series of group discussions on the psychology of sex. The remaining 30, tested at a somewhat later date, were "captive volunteers" from a psychology course who elected to participate in the group discussions on the psychology of sex in preference to several other experiments. Since the results obtained from these two samples were very similar, they were combined in the analysis presented here.

Each *S* was individually scheduled to "meet with a group." When she arrived at the experimental room, she was told by the experimenter (*E*) that he was conducting several group discussions on the psychology of sex. *E* informed her that she was joining a group that had been meeting for several weeks and that she was taking the place of a girl who had to leave the group because of scheduling difficulties. *E* stated that the discussion had just begun and that she would join the other members of the group after he had explained the nature of the experiment to her. The purpose of the foregoing instructions was to confront *S* with an ongoing group and thus make plausible the recorded discussion to which she was to be exposed.

E then "explained" the purpose of the experiment. He said that he was interested in investigating the "dynamics of the group discussion process." Sex was chosen as the topic for the groups to discuss in order to provide interesting subject matter so that volunteers for the discussion groups could be obtained without much difficulty. *E* continued as follows:

> But the fact that the discussions are concerned with sex has one major drawback. Although most people are interested in sex, they tend to be a little shy when it comes to discussing it. This is very bad from the point of view of the experiment; if one or two people in a group do not participate as much as they usually do in group discussions because they are embarrassed about sex, the picture we get of the group discussion process is distorted. Therefore, it is extremely important to arrange things so that the members of the discussion group can talk as freely and frankly as possible. We found that the major inhibiting factor in the discussions was the presence of the other people in the room. Somehow, it's easier to talk about embarrassing things if other people aren't staring at you. To get around this, we hit upon an idea which has proved very successful. Each member of the group is placed in a separate room, and the participants communicate through an intercom system using headphones and a microphone. In this way, we've helped people relax, and have succeeded in bringing about an increase in individual participation.

The foregoing explanation set the stage for the tape recording, which could now be presented to the *S* as a live discussion conducted by three people in separate rooms.

E then mentioned that, in spite of this precaution, occasionally some persons were still too embarrassed to engage in the discussions and had to be asked to withdraw from the discussion group. *S* was asked if she thought she could discuss sex freely. She invariably answered affirmatively. In the Control condition *S* was told, at this point, that she would be a member of the group.

In the other two conditions, *E* went on to say that it was difficult for him to ask people to leave the group once they had become members. Therefore, he had recently decided to screen new people before admitting them to the discussion groups. The screening device was described as an "embarrassment test" which consists of reading aloud some sexually oriented material in the presence of *E*. *S* was told that *E* would make a clinical judgment of her degree of embarrassment, based upon hesitation, blushing, etc. and would determine whether or not she would be capable of participating in the discussion group. He stressed that she was not obligated to take this test, but that she could not become a member unless she did. Only one *S* declined to take the test. She was excluded from the experiment. It was also emphasized, at this point, that the "embarrassment test" was a recent innovation and that the other members had joined the group before it was required for admission. These instructions were included in order to counteract any tendency to identify more strongly with the group as a result of feelings of having shared a common unpleasant experience. Such a process could conceivably bring about a greater preference for the discussion group on the part of *S*s in the Severe condition, introducing ambiguity in the interpretation of the results.

In the Severe condition, the "embarrassment test" consisted of having *S*s read aloud, from 3 × 5 cards, 12 obscene words, e.g., fuck, cock, and screw. *S*s also read aloud two vivid descriptions of sexual activity from contemporary novels. In the Mild condition, *S*s read aloud five words that were related to sex but not obscene, e.g., prostitute, virgin, and petting. In both the Severe and the Mild conditions, after each *S* finished reading the material, she was told that she had performed satisfactorily and was, therefore, a member of the group and could join the meeting that was now in progress.

It was of the utmost importance to prevent the *S* from attempting to participate in the discussion, for if she did, she would soon find that no one was responding to her statements and she would probably infer that the discussion was recorded. To insure their silence, all *S*s were told that, in preparation for each meeting, the group reads an assignment which serves as the focal point of the discussion; for this meeting, the group read parts of the book, *Sexual Behavior in Animals*. After

the *S* had indicated that she had never read this book, *E* told her that she would be at a disadvantage and would, consequently, not be able to participate as fully in this discussion as she would had she done the reading. He continued, "Because the presence of a participant who isn't contributing optimally would result in an inaccurate picture of the dynamics of the group discussion process, it would be best if you wouldn't participate at all today, so that we may get an undistorted picture of the dynamics of the other three members of this group. Meanwhile, you can simply listen to the discussion, and get an idea of how the group operates. For the next meeting, you can do the reading and join in the discussion." *S*s were invariably more than willing to comply with this suggestion. The above instructions not only prevented *S* from attempting to participate in the discussion but also served to orient her toward the actual content of discussion.

Under the guise of connecting the *S*'s headphones and microphone, *E* went into the next room and turned on the tape recorder. He then returned to the experimental room, put on the headphones, picked up the microphone, and pretended to break into the discussion which supposedly was in progress. After holding a brief conversation with the "members of the group," he introduced the *S* to them. Then he handed the headphones to her. The tape was timed so that at the precise moment that *S* donned her headphones, the "group members" introduced themselves and then continued their discussion.

The use of a tape recording presented all *S*s with an identical group experience. The recording was a discussion by three female undergraduates. It was deliberately designed to be as dull and banal as possible in order to maximize the dissonance of the *S*s in the Severe condition. The participants spoke dryly and haltingly on secondary sex behavior in the lower animals, "inadvertently" contradicted themselves and one another, mumbled several *non sequiturs*, started sentences that they never finished, hemmed, hawed, and in general conducted one of the most worthless and uninteresting discussions imaginable.

At the conclusion of the recording, *E* returned and explained that after each meeting every member of the group fills out a questionnaire expressing her reactions to the discussion. The questionnaire asked the *S* to rate the discussion and the group members of 14 different evaluative scales, e.g., dull–interesting, intelligent–unintelligent, by circling a number from 0 to 15. After completing the questionnaire, *S* made three additional ratings, orally, in response to questions from *E*. Nine of the scales concerned the *S*'s reactions to the discussion, while the other eight concerned her reactions to the participants.

At the close of the experiment, *E* engaged each *S* in conversation to determine whether or not she was suspicious of the procedure. Only one *S* entertained definite suspicions; her results were discarded.

Finally, the true nature of the experiment was explained in detail. None of the *S*s expressed any resentment or annoyance at having been misled. In fact, the majority were intrigued by the experiment and several returned at the end of the academic quarter to ascertain the results.

RESULTS AND DISCUSSION

The sum of the ratings for the 17 different scales provides an index of each *S*'s liking for the discussion group. The means and *SD*s for the three experimental conditions for this measure are presented in Table 1. Means and *SD*s are also presented in Table 1 separately for the eight scales which tapped the *S*s' attitudes toward the discussion and the seven scales which tapped their attitudes toward the participants. The significance of the differences between the means for the different conditions were determined by *t* tests. The *t* values and significance levels are presented in Table 2.

Examination of Table 1 shows that *S*s in the Severe condition rated both the discussion and the participants higher than did those in the Control and Mild conditions. The over-all difference between the ratings by *S*s in the

TABLE 1

MEANS OF THE SUM OF RATINGS FOR THE DIFFERENT EXPERIMENTAL CONDITIONS

Rating Scales	Experimental Conditions		
	Control ($N=21$)	Mild ($N=21$)	Severe ($N=21$)
Discussion [9]			
M	80.2	81.8	97.6
SD	13.2	21.0	16.6
Participants [8]			
M	89.9	89.3	97.7
SD	10.9	14.1	13.2
Total [17]			
M	166.7	171.1	195.3
SD	21.6	34.0	31.9

TABLE 2

SIGNIFICANCE LEVELS OF THE DIFFERENCES BETWEEN EXPERIMENTAL CONDITIONS

Rating Scales	Differences Between Conditions		
	Control-Severe	Mild-Severe	Control-Mild
Discussion [9]	$t = 3.66$ $P < .001$*	$t = 2.62$ $P < .02$	$t = .29$ N.S.
Participants [8]	$t = 2.03$ $P < .05$	$t = 1.97$ $P < .10$	$t = .15$ N.S.
Total [17]	$t = 3.32$ $P < .01$	$t = 2.33$ $P < .05$	$t = .49$ N.S.

* The *P* values given are based on both tails of the *t* distribution.

Severe condition and *S*s in the Control condition reaches the .01% level of significance. The over-all difference between the ratings by *S*s in the Severe initiation condition and *S*s in the Mild initiation condition reaches the .05 level.

These differences cannot be explained by differences in initial motivation to become members of the group, since *S*s (with varying degrees of motivation) were randomly assigned to the three experimental conditions. The differences in liking for the group must be considered a consequence of the unpleasant experience. The results clearly substantiate the hypothesis: persons who undergo a severe initiation to attain membership in a group increase their liking for the group. This hypothesis follows directly from Festinger's theory of cognitive dissonance. According to the theory, *S*s in the Severe initiation condition held the cognition that they had undergone a painful experience to become members of the discussion group. Then they listened to a dull, banal discussion. Negative cognitions about the discussion which they formed from listening to it were dissonant with the cognition that they had undergone a painful experience to gain membership in this group. The presence of dissonance leads to pressures to reduce it. *S*s in this condition could reduce their dissonance either by denying the severity of the initiation or by distorting their cognitions concerning the group discussion in a positive direction. The initiation of the *S*s in the Severe condition was apparently too painful for them to deny — hence, they reduced their dissonance by overestimating the attractiveness of the group.

There was no appreciable difference between the ratings made by *S*s in the Control condition and those made by *S*s in the Mild condition. It would seem that the Mild condition was so devoid of unpleasantness as to constitute little investment in the group. Hence, little dissonance was created. If any dissonance did occur in this situation it would be more realistic for the *S* to reduce it by minimizing the pain of the initiation, than by distorting her cognitions concerning the discussion. Thus, it is not an initiation per se that leads to increase in liking for a group. The initiation must be severe enough to constitute a genuine investment and to render it difficult to reduce dissonance by playing down the extent of the pain involved.

An examination of Table 1 shows that the rating scales concerning the discussion show greater differences between the conditions than the scales dealing with the evaluations of the participants in the discussion. There are at least two possible explanations for this result: (*a*) It may be easier for people to express negative criticism about an impersonal discussion than about the people involved. Thus, *S*s in the Control and Mild conditions may have inflated their ratings of the participants to avoid making negative statements about fellow college students. (*b*) It is possible that *S*s in the Severe condition had less need to distort their perception of the participants than of the discussion itself. The dissonance of the *S*s in the Severe condition resulted from the actual discussion: they experienced dissonance between going through an unpleasant experience and taking part in worthless uninteresting discussions. The most direct way for them to reduce this dissonance would be to change their perceptions of the discussion in a positive direction. The participants in the discussion were peripheral to the cause of dissonance. If *S*s in the Severe condition had less need to distort their perceptions of the participants than their perception of the discussion, their evaluations of the participants could be expected to be closer to the evaluations of the participants made by *S*s in the Control and Mild conditions.

SUMMARY AND CONCLUSIONS

An experiment was conducted to test the hypothesis that persons who undergo an unpleasant initiation to become members of a group increase their liking for the group; that is, they find the group more attractive than do persons who become members without going through a severe initiation. This hypothesis was derived from Festinger's theory of cognitive dissonance.

College women who volunteered to participate in discussion groups were randomly assigned to one of three experimental conditions: A *Severe* initiation condition, a *Mild* initiation condition, and a *Control* condition. In the Severe condition, subjects were required to read some embarrassing material before joining the group; in the Mild condition the material they read in order to join the group

was not very embarrassing; in the Control condition, subjects were not required to read any material before becoming group members. Each subject listened to a recording that appeared to be an ongoing discussion being conducted by the group which she had just joined. Afterwards, subjects filled out a questionnaire evaluating the discussion and the participants. The results clearly verified the hypothesis. Subjects who underwent a severe initiation perceived the group as being significantly more attractive than did those who underwent a mild initiation or no initiation. There was no appreciable difference between ratings by subjects who underwent a Mild initiation and those by subjects who underwent no initiation.

REFERENCE

FESTINGER, L. *A theory of cognitive dissonance.* Evanston: Row, Peterson, 1957.

Received June 9, 1958.

ANTI-SEMITISM AND THE DISPLACEMENT OF AGGRESSION[1]

LEONARD BERKOWITZ
University of Wisconsin

As is well known, the psychoanalytic principle of displacement maintains that aggressive tendencies denied expression against the object instigating the aggression tend to be directed against noninstigating objects. Long part of the social science folklore, this principle is given empirical support in three studies reported by Dollard, Doob, Miller, Mowrer, and Sears (1939, pp. 42–44) in their now classic work, *Frustration and Aggression*.

In all of these experiments the frustrated organisms displayed increased unfriendliness toward objects *that happened to be available*. However, negative findings obtained in more recent research (e.g. Stagner & Congdon, 1955) indicate that the choice of a target for displaced aggression is not necessarily solely a function of availability. As Miller (1948) has shown, an object's likelihood of receiving displaced aggression depends upon such stimulus factors as its similarity to the instigator and whether the direct overt hostility against the instigator is prevented by the instigator's absence or by internal conflict. In the former case, according to Miller's analysis, displaced responses occur to other similar objects, and the strongest aggressive act occurs to the most similar object present. If there is conflict, however, on the assumption that the gradient of generalization of the interfering responses is steeper than that of the aggressive responses they inhibit, the strongest hostile response is predicted to occur to stimulus objects that have an intermediate degree of similarity to the original instigating object. Supporting evidence is found in two recent correlational investigations (Murney, 1955; Wright, 1954) employing fantasy productions.

The objective stimulus situation may not be the only determinant of the tendency to displace aggression, however. Many students of intergroup relations have suggested that variations in this tendency also are associated with the personality characteristics of the angered individuals. Thus, according to the "scapegoat" theory of prejudice, highly prejudiced individuals are more likely than those who are less prejudiced to reduce their hostility by displacing it upon minority groups.[2] But although this contention has had a considerable vogue among social scientists (e.g. Brown 1942; Fenichel, 1946), there have been few attempts to test the validity of this theory experimentally, and at least one test (Lindzey, 1950) has yielded negative results. Nor have there been many attempts to spell out explicitly the processes accounting for the scapegoat reaction. For example, if the prejudiced person's hostility toward minority groups is a case of displaced aggression, is this due solely to a high level of aggressive drive or to others factors unique to his type of personality, such as his learning to regard others (particularly minority group members) as just as frustrating as the present instigator of his hostility? Fromm (1941) and Maslow (1943) are in agreement in characterizing the authoritarian personality, who, of course, is likely to be highly prejudiced, as possessing a relatively high level of hostility, while Siegel (1956) has found that the measure of this personality type, the F scale, is positively correlated with an index of manifest hostility. If the former alternative were correct, then, there should be a greater incidence of displacement in experimentally angered than in less angered subjects (*S*s), regardless of whether these *S*s are characteristically prejudiced or not, as long as their level of aggression is the same. The studies by Miller and Bugelski (Dollard, Doob, Miller, Mowrer, & Sears, 1939, pp. 42–44) sup-

[1] This study was supported by Research Grant M1540 from the National Institute of Mental Health, U.S. Public Health Service. Douglas Holmes rendered invaluable assistance in collecting the present data.

[2] It is not clear from most descriptions of the role of displaced aggression in social prejudice as to whether the displacement is a general reaction or is elicited only by objects having certain stimulus characteristics, e.g., Negroes or Jews. At least one paper in this area (Miller & Bugelski, cited in Dollard et al., 1939) suggests that the former is a likely possibility.

port this possibility. However, if other individual characteristics affect the displacement tendency (as is suggested by the failures to replicate the Miller-Bugelski results) and if these are correlated with persistent intolerance toward minority groups, prejudiced *Ss* should be more likely to respond to a frustration with displaced aggression than less prejudiced individuals, even though both groups are equally angry.

The present study was designed to test these alternatives. *Ss* differing in their level of anti-Semitism were either angered or not by *E* and then assigned to work with a partner on a neutral task. To the extent that the aroused anti-Semitic *Ss* exhibit more hostility toward their partners than the similarly angered less prejudiced *Ss*, reliable individual characteristics, as well as the objective nature of the stimulus objects, will have been shown to affect the displacement of aggression.

METHOD

Subjects

The *Ss* were 48 female volunteers from introductory psychology classes at the University of Wisconsin distributed evenly among the eight conditions in a 2 × 2 × 2 factorial design. The girls were classified as either high or low in anti-Semitism and, within each of these groups, either high or low in aggressive drive, on the basis of their responses to a questionnaire administered about a month before the experiment. The former classification made use of the *Ss*' scores on 17 Likert-type items comprising a modified version of the anti-Semitism scale developed by the California investigators of the authoritarian personality (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950) with the high and low anti-Semitic *Ss* (High A-S and Low A-S) taken from the upper and lower thirds of the total distribution of scores on this scale. The aggressive drive classification was an attempt to control for the level of hostility within the *Ss*, employing their scores on Siegel's Manifest Hostility Scale (MHS)[1] (1956), much as the Taylor Manifest Anxiety Scale has been used in controlling the level of anxiety drive. However, as we would expect on the basis of Siegel's results, it was found that anti-Semitism was significantly associated with manifest hostility level. Therefore, while both the high and low A-S *Ss* were dichotomized into high and low MHS *Ss*, the high A-S *Ss* as a group are higher in manifest hostility than the low A-S *Ss* ($p < .05$).

[1] More recent evidence points to important difficulties in the interpretation of scores on this scale. Two studies (Berkowitz, in press; Hokanson & Gordon, 1958) agree in reporting a negative correlation between MHS scores and the increase in intensity of aggressive behavior after arousal.

TABLE 1
MEAN SCORES IN EACH OF THE EXPERIMENTAL CONDITIONS[a]

MHS Level Measure	Hostility Arousal		Non-Arousal	
	High A-S	Low A-S	High A-S	Low A-S
High				
MHS	27.2	22.5	31.2	23.3
Annoyance[b]	13.5	12.0	7.5	6.0
Like P[c]	18.8	13.5	12.8	17.7
Low				
MHS	16.5	9.7	13.7	9.3
Annoyance	11.5	6.7	8.0	7.7
Like P	18.0	10.8	15.5	15.0
All				
MHS	21.8	16.1	22.4	16.3
Annoyance	12.5	9.3	7.8	6.8
Like P	18.4	12.2	14.2	16.3

[a] $N = 6$ in each condition.
[b] The higher the score the less *S* enjoyed the experiment.
[c] The higher the score the less *S* liked her partner.

Table 1 presents the mean MHS score in each of the experimental conditions.

Procedure

The experimental manipulations began as soon as any given *S* showed up for her appointment, ostensibly for an experiment on "problem-solving." In the *Hostility Arousal* condition, *E* met *S* with the statement that she was late and sarcastically asked whether this was always true. He then took her into another room (after first closing a door in her face) where he introduced *S* to a concept formation task. *E* deprecated *S*'s performance during this 10-minute task and questioned her ability to do well in college. In the *Non-Arousal* condition, *E* maintained either a neutral or moderately friendly attitude toward *S*. Following the completion of the concept formation task, *E* brought another girl into the room saying that this girl also had just finished the first phase of the experiment. In actuality, the new girl was a paid confederate of *E*. They were now to work together on another kind of problem: to decide how a mediation board would have voted in a supposedly real labor-management dispute. The two were given information pertaining to the history of the dispute and the lives of the mediation board members. After 15 minutes of discussion the *Ss* had to reach some sort of decision. *E* never told the confederate whether her partner was "aroused" or not or what her level of anti-Semitism was. Since it has been shown that anti-Semitic *Ss* are more likely to guess that an ambiguous person is Jewish than less prejudiced *Ss* (Elliott & Wittenberg, 1955), the confederate was given a Jewish name (Miss Cohen) in half of the cases within each condition and a Gentile name (Miss Johnson) in the remaining cases. Inspection of the data indicates that no significant effects can be traced to the labeling of the confederate as Jewish or not.

Dependent Variables

Following the discussion, the *S*s were asked to complete a brief questionnaire indicating their attitudes toward the experiment. One of the items in this questionnaire was designed to determine whether the manipulation of hostility arousal had been successful. The question obviously could not ask whether *E* had angered *S*, but it was assumed that if he had, *S* would experience this as a frustration and therefore would be less likely to have enjoyed the experiment. The question therefore read simply: "Did you enjoy the experiment?" with the answers ranging from "Definitely yes" (score = 0) to "Definitely no" (score = 23). This score is referred to below as the Annoyance measure.

The displaced aggression measure consists of two correlated items assessing *S*'s liking for her partner. The items are as follows: "If you were assigned to another two-person task, to what extent would you want to work with your present partner again?" and, "If it involved any effort on your part, would you want to get to know your present partner any better?" The answers again ranged from "Definitely yes" (= 0) to "Definitely no" (= 23), with the total liking-for-partner score (Like P) consisting of the sum of the responses to the two items.

At the completion of the experiment the "aroused" *S*s were informed that they had been deliberately angered "in order to study the effects of stress upon problem-solving and relationships with others." Most of the *S*s expressed an interest in the experiment, and all promised not to discuss the nature of the deceptions with others.

RESULTS

The only significant effect with the Annoyance measure, by analysis of variance, was the main effect for the arousal conditions ($F = 7.09$; $p = .02$). As can be seen in Table 1, the *S*s in the Hostility Arousal condition enjoyed the experiment less than the *S*s in the Non-Arousal condition, indicating that the experimental manipulation was effective. Contrary to Lindzey's (1950) finding that anti-Semitic *S*s are more disturbed by a frustration than less anti-Semitic *S*s, the High A-S *S*s were not significantly more annoyed by the experiment than the less prejudiced *S*s ($F = 2.25$; $p > .10$).

Table 2, reporting the results of the analysis of variance of the Like *P* data, provides support for the hypothesis that prejudiced *S*s are more likely to displace aggression when angered than less prejudiced *S*s. As is shown in the first table, the High A-S *S*s in the Non-Arousal condition, though possessing a significantly higher level of manifest hostility, displayed as much liking for those working with them as the Low A-S *S*s.[4] However, when aggression was instigated by a third party, the High A-S *S*s tended to increase in hostility toward their partners, while the Low A-S *S*s became somewhat friendlier. The former *S*s apparently displaced their aggression upon an available, and perhaps safe, object. The interaction of hostility arousal with anti-Semitism is significant at less than the .05 level of confidence.

TABLE 2
ANALYSIS OF VARIANCE OF THE LIKING-FOR-PARTNER SCORES

Source	*df*	Mean Squares	*F*	*P*
1. Arousal conditions	1	0.02		
2. Manifest hostility	1	9.19		
3. Anti-Semitism	1	50.02		
4. 1 × 2	1	9.19		
5. 1 × 3	1	212.52	5.95	< .05
6. 2 × 3	1	38.52		
7. 1 × 2 × 3	1	9.19		
8. Residual	40	35.72		

One conceivable explanation for this displacement can be found in the theoretical formulations presented by Adorno et al, (1950).They suggest that the authoritarian (or, in our case, the prejudiced person) tends to project hostility upon others in his environment in a diffuse manner. In the present situation, this diffuse projectivity implies that all others are blamed for the annoyance felt by *S*. The prejudiced person thus denies that the frustration he is now suffering is due to his own actions, and broadly categorizes most or all others in his environment as frustration-producers, as the people responsible for his present plight. Consistent with this interpretation, Lesser (1958) has demonstrated that anti-Semitic *S*s characteristically manifest extrapunitive reactions to frustrations. They attack others in denying their responsibility for the annoyance.

The last item in the postexperiment questionnaire can be used to determine whether

[4] In a recent experiment (Berkowitz, in press), it was found that MHS scores were positively correlated with the intensity of the hostility displayed in *S*s' first impressions of others, but that after communicating with these others the *S*s low in MHS increased in this unfriendliness to the level exhibited by High MHS *S*s.

TABLE 3
MEAN SCORES ON ITEM, "DO YOU FEEL THAT THIS TASK IS A GOOD MEASURE OF YOUR PROBLEM-SOLVING ABILITY?"

MHS Level	Hostility Arousal		Non-Arousal	
	High A-S	Low A-S	High A-S	Low A-S
High	15.5[a]	10.5	12.0	10.0
Low	15.5	10.3	10.8	12.7
All	15.5	10.4	11.4	11.3

Source	Analysis of Variance		F	P
	df	Mean Squares		
1. Arousal conditions	1	30.09		
2. Manifest hostility	1	1.34		
3. Anti-Semitism	1	80.09		
4. 1 × 2	1	2.07		
5. 1 × 3	1	74.99	4.04[b]	.06
6. 2 × 3	1	10.07		
7. 1 × 2 × 3	1	12.02		
8. Residual	40	18.55		

[a] The higher the score the more strongly *S* denied that the task was a good measure of her problem-solving ability.
[b] $F_{.05} = 4.08$.

the High A-S's diffuse projectivity tendency could account for the displacement of aggression. This item reads, "Do you feel that this task is a good measure of your problem-solving ability?" The *S* who answers in the No direction may be exhibiting a projectivity reaction. That is, in saying that his performance is not a reflection of his true ability he denies that this particular performance is associated with his self.

The results of the analyses with this measure are given in Table 3. A comparison of this table with Table 1 demonstrates that answers to this last item are not a simple reflection of *Ss*' annoyance with the experiment; the results do not parallel the results with the previously discussed Annoyance measure. For example, the "aroused" *Ss* generally are *not* more likely to say that the task is a poor measure of their problem-solving ability, as would be expected if this statement were solely an act of aggression toward *E*.

There is somewhat of a parallel, however, between the present results and the findings with the Like P measure, particularly with regard to the interaction of hostility arousal and anti-Semitism. As was the case with the expressions of liking for their partner, there is no difference between the High A-S and Low A-S *Ss* in the Non-Arousal condition, while the differences do become significant in the "aroused" condition. The High A-S *Ss* presumably denied real responsibility for their problem-solving actions, but apparently only when subjected to stressful conditions.

DISCUSSION

The present results demonstrate that the likelihood that displaced aggression will be directed toward an object is a function of the angered person as well as of objective stimulus factors. As psychoanalytically oriented social scientists have long maintained, highly prejudiced people are more likely to displace hostility diffusely upon their environment than less prejudiced individuals. (It is important to emphasize, however, that this displacement is manifested only when anger is aroused.) By comparison, the less prejudiced, though almost equally annoyed by the instigation, apparently channelize their aggression more effectively. While the evidence is far from unequivocal, these findings also suggest that displacement is not solely dependent upon the strength of the aggressive drive. Additional factors must be introduced. One possibility is that the High A-S *Ss*, when angered, were more likely than the Low A-S *Ss* to inhibit direct (though covert) acts of hostility toward the higher status *E* and, therefore, had to release this aggression onto available others. Another possibility for which the present study has some support has to do with the prejudiced individual's tendency to exhibit a diffuse projectivity, i.e., to generally blame others for the annoyance she feels. According to this interpretation, the angered High A-S *S* broadly categorizes all or most others, including her co-workers, as responsible for her unpleasant plight and denies self-responsibility for her frustrations. The resulting aggression is then directed to these others. Another manifestation of this presumed denial of self-responsibility demonstrated in the present experiment is the relatively strong tendency displayed by the "aroused" High A-S *Ss* to state that their present performance did not reflect

their true problem-solving ability. In this case, they are saying that their behavior is not truly associated with their self.

This positive support for the scapegoat theory of prejudice must be contrasted with Lindzey's (1950) negative results. In all probability, this difference stems largely from the different measures of displaced aggression in the two experiments. Lindzey used two indices of displacement, one based upon the number of incidents in the *S*s' TAT protocols in which "self" figures carried out aggressive behavior against "nonself" figures, the other the extrapunitive scale of the P-F Study. Conceivably, these measures could have low reliability and validity as indices of displacement. It also is reasonable to assume that the TAT index assesses, not displaced but, rather, a general fantasy aggression. Support for this possibility is found in a recent experiment by Hokanson and Gordon (1958). They noted that *S*s with low scores on Siegel's MHS tended to increase in fantasy aggression after arousal, while the High MHS *S*s had lower scores after arousal. Since anti-Semitism is positively correlated with MHS, this finding is consistent with Lindzey's observation that the less anti-Semitic *S*s had a somewhat higher fantasy aggression score (our term for his measure) after arousal than the more highly prejudiced *S*s.

There is little doubt either that the displacement of aggression occurs or that this phenomenon can play an important part in many (but certainly not all) cases of social prejudice. However, it is equally apparent that displacement is not necessarily elicited in angered persons simply by preventing attacks against the instigator of their hostility. A target's likelihood of receiving displaced hostility varies depending upon a number of factors, such as (*a*) whether the direct attacks against the instigator of aggression are prevented by internal conflict or by his absence, (*b*) the degree of similarity between the given available target and the perceived instigator of aggression, (*c*) the relative strength of the tendencies to aggress and to inhibit the aggression, (*d*) the nature of the threat arousing the anger, and (*e*) the personality of the angered individual. Taking these in order, Miller's (1948) theoretical analyses and the research stimulated by him (e.g., (Murney, 1955; Wright, 1954)) demonstrate the importance of the first three factors. Feshbach and Singer (1957) show that a personal or ego threat often results in increased hostility and in social prejudice (i.e., displaced hostility), while a "shared' threat results in a decrement in the expression of social prejudice, and the present study indicates that characteristically prejudiced (anti-Semitic) *S*s are more likely to displace aggression when angered than less prejudiced *S*s.

SUMMARY

The present study was designed to demonstrate that an object's likelihood of receiving displaced aggression varies not only with the objective stimulus factors discussed by Miller (1948) and others (Dollard et al., 1939) but also with the personality characteristics of the angered person. In particular, it was hypothesized, in accord with some versions of the scapegoat theory of prejudice, that highly anti-Semitic *S*s would be more likely to displace aggression when angered than less anti-Semitic *S*s. The results support the hypothesis. Several factors possibly accounting for this displacement are discussed. Evidence is provided consistent with the view that prejudiced people are likely to exhibit a diffuse projectivity when frustrated, i.e., to generally blame others for the annoyance they feel and deny self-responsibility for their plight.

REFERENCES

ADORNO, T., FRENKEL-BRUNSWIK, E., LEVINSON, D., & STANFORD, R. *The authoritarian personality*. New York: Harper, 1950.

BERKOWITZ, L. Manifest hostility level and hostile behavior. *J. soc. Psychol.*, in press.

BROWN, J. F. The origin of the anti-Semitic attitude. In I. Graeber & S. H. Britt (Eds.), *Jews in a gentile world*. New York: Macmillan, 1942.

DOLLARD, J., DOOB, L., MILLER, N., MOWRER, O., & SEARS, R. *Frustration and aggression*. New Haven: Yale Univer. Press, 1939.

ELLIOTT, D., & WITTENBERG, B. Accuracy of identification of Jewish and non-Jewish photographs. *J. abnorm. soc. Psychol.*, 1955, **51**, 339–341.

FENICHEL, O. Elements of a psychoanalytic theory of anti-Semitism. In E. Simmel (Ed.), *Anti-Semitism: A social disease*. New York: International Universities Press, 1946.

FESHBACH, S., & SINGER, R. The effects of personal and shared threats upon social prejudice. *J. abnorm. soc. Psychol.*, 1957, **54**, 411–416.

FROMM, E. *Escape from freedom*. New York: Rinehart, 1941.

HOKANSON, J., & GORDON, J. The expression and inhibition of hostility in imaginative and overt behavior. *J. abnorm. soc. Psychol.*, 1958, **57**, 327–333.

LESSER, G. S. Extrapunitiveness and ethnic attitude. *J. abnorm. soc. Psychol.*, 1958, **56**, 281–282.

LINDZEY, G. An experimental examination of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.*, 1950, **45**, 296–309.

MASLOW, A. Authoritarian character structure. *J. soc. Psychol.*, 1943, **18**, 401–411.

MILLER, N. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.

MURNEY, R. *An application of the principle of stimulus generalization to the prediction of object displacement.* Washington, D. C.: Catholic Univer. America Press, 1955.

SIEGEL, S. The relationship of hostility to authoritarianism. *J. abnorm. soc. Psychol.*, 1956, **52**, 368–372.

STAGNER, R., & CONGDON, C. Another failure to demonstrate displacement of aggression. *J. abnorm. soc. Psychol.*, 1955, **51**, 695–696.

WRIGHT, G. Projection and displacement: A cross-cultural study of folktale aggression. *J. abnorm. soc. Psychol.*, 1954, **49**, 523–528.

Received June 12, 1958.

A TEST OF WITKIN'S FIELD-DEPENDENCE HYPOTHESIS[1]

HARL H. YOUNG

Veterans Administration Hospital, Denver, Colorado

THE ten-year study of personality and perception by Witkin and his colleagues (Witkin, Lewis, Hertzman, Machover, Meissner, & Wapner, 1954) grew out of attempts to determine whether visual or kinesthetic cues are dominant in the process of orienting one's self to the vertical in space. In studying this process, wide individual differences were observed in scores on the orientation tasks. Some people were found to rely mainly on the visual field, while others used sensations resulting from the pull of gravity on their bodies as reference frames in establishing the upright. The investigators labeled the tendency to rely on the visual field as "field-dependence," while the inclination to use body cues, which apparently requires an analytic approach, was called "field-independence." These two terms were conceived as representing the extremes of a continuum that reflect a dimension of personality roughly corresponding to activity-passivity in dealing with the environment.

In order to test this general hypothesis, the Witkin group administered the orientation tasks to a large group of subjects. Measures based on these perceptual situations were then compared with measures of personality based on several widely used clinical techniques, including the assessment interview, Rorschach, TAT, and the Draw-a-Person test. The investigators interpreted their results as indicating considerable validity to the general field-dependence hypothesis.

However, critical reviews of this work (e.g., Holtzman, 1955) have called attention to some methodological weaknesses. One of the most serious of these involved the authors' almost complete lack of concern for the reliability of their personality measures. No estimates of interscorer agreement were presented, nor were data reported concerning intratest or test-retest consistency. A second criticism was that the correlations with perceptual scores based on the personality indices were spurious to an unknown degree because the protocols were first examined for pertinent variables and then placed back in the sample. Although some cross-validation data are presented, these were obtained on a very different sample (hospital patients) and are not entirely satisfactory for validating results on normal college students.

These considerations make a replication highly desirable, to provide a more stringent test of the general field-dependence hypothesis. This was the objective of the present research. Specific hypotheses tested in the present study were as follows:

Hypothesis I. *There is a common factor of field-dependence running through all the perceptual tests used in this study.* Successful performance on each perceptual test requires the separation of an item from its context or surround. The greater the subject's performance is retarded by the surround, the more field-dependent he is said to be. If these perceptual measures actually are highly intercorrelated, they can then be used as criterion variables for evaluation of the hypotheses relating personality to perception which follow.

Hypothesis II. *Self attitudes of passivity, dependence, distrust of one's own feelings, and bodily experiences as revealed projectively in drawings of humans are significantly related to field-dependent perceptual performance.*

Hypothesis III. *Responses to inkblot materials suggesting a lack of effectiveness in coping with environmental demands and one's own needs are significantly related to perceptual performance*

Hypothesis IV. *Responses to inkblots im-*

[1] This report is adapted from a dissertation submitted to the graduate faculty of the University of Texas in partial fulfillment of the requirements for the degree of doctor of philosophy. The writer wishes to express his appreciation to Wayne H. Holtzman, chairman of the thesis committee, for his generous assistance in connection with the study. Gratitude is also expressed to the other members of the committee, Karl M. Dallenbach, Philip Worchel, Harry Helson, Ira Iscoe, Royal Embree, and O. B. Douglas. Thanks are extended to David Proctor for his assistance with scoring and reliability problems. The research was facilitated by a grant from the Faculty Research Fund of the University of Oklahoma while the author was on the faculty there.

plying a lack of introspectiveness, an impoverished inner life, and self-distrust are related to field-dependent perceptual performance.

Hypothesis V. *The greater the discrepancy between one's Self ratings and one's Ideal-Self ratings, the more field-dependent behavior is exhibited.*

Hypothesis VI. *The tendency to depreciate one's self is associated with field-dependent perceptual performance.*

Hypotheses I through IV follow directly from Witkin's work, while the final two hypotheses represent extensions of the general field-dependence hypothesis. The Witkin (1954) group summarizes the characteristics of field-dependent and field-independent persons as follows:

> . . . Field-dependent persons tend to be characterized by passivity in dealing with the environment; by unfamiliarity with and fear of their own impulses, together with poor control over them; by lack of self-esteem; and by the possession of a relatively primitive, undifferentiated body image. Independent or analytical perceptual performers, in contrast, tend to be characterized by activity and independence in relation to the environment; by closer communication with, and better control of, their own impulses; and by relatively high self-esteem and a more differentiated, mature body image (p. 469).

Hypotheses V and VI regarding self-ideal discrepancy and self-depreciation follow from this statement.

METHOD

Subjects

Ninety-seven undergraduates (49 males, 48 females) from a variety of departments at the University of Texas served as *S*s during the summer session of 1956. This group was deliberately chosen because of its greater heterogeneity than the usual college population.

Procedure

Six tests, requiring a total testing time of 3 to 4 hrs., were administered to each *S*.

Perceptual tests. Three tests were administered: Witkin's Rod-and-Frame test (1954), Jackson's short form of the Embedded Figures test (1956), and the Chair-Window test (Barrat, 1955). Each of these will be described briefly.

The Rod-and-Frame test requires that the *S*, while in a darkened room, adjust to the true (gravitational) vertical a luminous rod mounted within a luminous square frame. The frame and the chair in which the *S* sits may each be tilted 28 degrees in either direction while the *S* makes each judgment. A protractor mounted on the back of the apparatus permits direct readings of the amount of error in degrees. The apparatus is such that large error scores indicate greater reliance on the field elements provided by the frame. Three scores are derived from the test. Series 1 contains trials in which the body and frame are tilted in the same direction. Series 2 trials are conducted with the body and frame in opposite directions, and Series 3 represents trials with the body in the upright position.

Jackson's short form of the Embedded Figures test is a briefer version of Witkin's modification of the Gottschaldt Figures. There are 12 items, each of which requires the *S* to isolate a small, simple geometrical figure in a more complex figure without directly comparing the two. The score is the mean time per item.

Barrat and Fruchter's Chair-Window test requires that *S* determine through which "window" of a transparent "room" a chair would have to be viewed in order to look like given pictures of chairs in various perspectives. There are 32 such items, and the score is the number correct. Since the other perceptual tests involve time and error scores they were expected to correlate negatively with the Chair-Window. Adequate reliability estimates for each of the tests have been reported by the authors. This test was not a part of the original field-dependence battery but was included in the present study because successful performance on it apparently requires the same type of "analytic" approach thought to be critically important in the Rod-and-Frame and Embedded Figures tests.

Personality tests. Those used in this study were Machover's Draw-a-Person test (1954), an experimental form of a new inkblot test developed by Holtzman (1956, 1958; Young, 1958), and Worchel's Self-Activity Inventory (1956).

Draw-a-Person test scores were based on a modification (Bauman, 1951) of the original Machover scales presented by Witkin et al. (1954). Each *S*'s score is presumed to be a measure of lack of trust of his own feelings and body experiences and hence to tap field-dependence. Interscorer reliability based on independent scorings of 10 protocols drawn at random from the present sample was .82.

The Holtzman Inkblot test represents an attempt to obviate the well-known deficiencies of the Rorschach as a research instrument. Major innovations include increasing the number of inkblots to 40, while limiting responses to one per card, and extending the variety of stimulus colors, patterns, and shadings used in the original Rorschach materials. Statistically parallel forms have been prepared. Both content and interscorer reliability estimates are well established (Holtzman 1956, 1958; Young, 1958).

Two summary scores derived from this instrument were patterned after the WPC or "coping" score and the MHFA or "introspective" score developed by Hertzman for use with the Rorschach in the Witkin study (1954, pp. 207–215). Hertzman scored the protocols in the usual manner and then assigned each of the W, P, or C scores and the M, H, F, and A scores in terms of specially developed criteria. In each case, a score of one was given if the kind of response was scarce or absent. For example, an H score of one was assigned if Hd/H + Hd equalled .5 or greater. These

separate scores were then summed for each *S* across WPC to obtain a measure of *lack* of ability to cope with environmental stresses. The MHFA score, intended as an index of lack of introspectiveness and awareness, was obtained by summing MHF and then subtracting A.

Criteria for the assignment of these scores in the present study were identical, in many cases, with those used by Hertzman (i.e., C, H, and A). However, some adaptations were necessary because of differences in P and the number of responses obtained using the different methods. In the present study, a percept given to the same area of the blot by 18% or more of the population was designated a "popular." This value was chosen because it was very near the median percentage of individuals reporting identical percepts to each of the 40 inkblots. Hertzman's criterion for the assignment of the P score is five or fewer populars. Similarly, several other criteria (W, M, and F) were stated in terms of frequencies. Hence, in order to compensate for differences in number of responses, the frequencies involved in Hertzman's criteria were simply doubled for use as criteria in the present study on the assumption that college students tend to give in the neighborhood of 20 responses to the Rorschach, while about 40 responses is usual for the Holtzman Inkblot test.

Although the Witkin group employed the conventional individual Rorschach examination, the Holtzman Inkblot test was given by means of projected slides and test booklets in small groups, since pilot studies had shown only minor differences between the distributions of scores obtained with group versus individual administrations of the Holtzman as well as between the Rorschach scoring distributions and the Holtzman variables. In view of the closely similar distributions obtained, it seemed justifiable to employ the Rorschach scoring categories on the Holtzman. The prime reason for using the Holtzman was that the new test in contrast with the Rorschach permits dependable reliability estimates of the scoring variables (Holtzman 1958).

Worchel's Self-Activity Inventory contains a series of items with negative connotations. The *S* is required to rate himself, his ideal self, and his conception of the average person on a five-point scale for each item. The summed Self ratings are therefore an index of self-depreciation. The summed absolute discrepancies between Self and Ideal ratings yield a measure of the degree to which a person's conception of himself differs from his idealized self. Worchel (1956) has reported satisfactory test-retest reliability estimates.

The three personality tests chosen for use in this study may be viewed as reflecting differing levels of the personality structure. The self-concept variables are simply conscious evaluations of the *S*, the Human Figure drawings may be assumed to be less influenced by conscious control, while it is widely believed that responses to inkblots reflect primarily unconscious material, as well as ways of coping with new and unstructured stimuli.

Scores on the 10 variables were derived from the foregoing tests and intercorrelated separately for each sex. The procedure yielded two 10 × 10 matrices, one for each sex.

RESULTS

The results are presented in Table 1.

Hypothesis I held that individuals showing field-dependent behavior on one of the perceptual tests would display similar behavior on the others. Inspection of the intercorrelations among the three series of the Rod-and-Frame, the Embedded Figures, and the Chair-Window indicates that 18 of the 20 coefficients are significant at the .05 level or better. The other two correlations just missed significance. It is clear that these tests are measuring a common factor to a considerable degree. Thus, Hypothesis I is confirmed.

Special attention is called to the correlations involving the Chair-Window test and the other perceptual measures. In every case the values are statistically significant at the .05 level. Moreover, the test apparently measures field-dependence more consistently for both sexes than does the Embedded Figures test.

Hypothesis II states that self attitudes of passivity, dependence, distrust of one's own feelings and bodily experiences, as revealed in the Figure drawings, are related to field-dependent perceptual performance. Inspection of Table 1 shows that two of 10 correlations are statistically significant. Series 2 of the Rod-and-Frame and the Embedded Figures scores are each significantly related to the DAP score for men. The Embedded Figures correlation is .52, which is highly significant, and the Rod-and Frame Series 2 value is .31. It is interesting to note that the correlation values are consistently higher for men than women. Hypothesis II is thus partially confirmed for men and not confirmed for women.

Eight of the above 10 correlations represent comparisons identical with those made by Witkin et al. The former investigators also found consistently higher values for men, but in their study *all* eight values were significant at the .01 level.

Hypothesis III stated that responses to inkblots suggesting a lack of effectiveness in coping with environmental demands (the WPC score) are significantly related to field-dependent perceptual performance. Table 1 indicates that two of five correlations are significant for women, while 1 of 5 values is significant for men. Among women, Series 3

TABLE 1
INTERCORRELATIONS AMONG MAJOR VARIABLES

Variable	RFT$_1$	RFT$_2$	RFT$_3$	EFT	CW	DAP	WPC	MHFA	SI	S
	(Males, N = 49)									
RFT Series 1		84	44	26	−30	21	−06	32	16	03
RFT Series 2	61		42	24	−42	31	13	22	06	00
RFT Series 3	57	63		37	−42	19	−04	44	24	−05
Embedded Figures test	65	50	38		−47	52	27	−03	14	01
Chair-Window	−35	−33	−48	−49		−26	−29	−22	−08	08
Draw-a-Person	09	06	22	16	−22		06	03	−07	−25
Coping (WPC)	24	21	43	25	−36	08		14	01	−01
Introspectiveness (MHFA)	19	24	30	04	−15	−06	23		21	−17
Self-Ideal Discrepancy	10	−16	−14	03	01	−39	−17	−21		55
Self Score	−04	−15	−11	−01	00	−13	−17	−23	68	
	(Females, N = 48)									

Note.—An r of .29 is significantly different from zero at the .05 level, while an r of .37 is significantly different from zero at the .01 level of confidence. All decimal points omitted.

of the Rod-and Frame is significantly related to the WPC or coping score (.43), but not so for men (−.04). For men, the WPC score is significantly related to the Chair-Window (−.29). Hypothesis III is thus partially confirmed for both sexes. In comparing these data with Witkin's results it is seen that *all* eight corresponding values reported by Witkin are statistically significant. The original study as well as the present research found these relationships to be stronger for women.

Hypothesis IV held that responses to inkblots implying a lack of introspectiveness (MHFA) are positively related to field-dependent perceptual performance. The table reveals that two of five correlations are statistically significant among males, while one of five coefficients is significant for females. Specifically, MHFA correlates .32 and .44 with Series 1 and 3 of the Rod-and Frame for men, while the MHFA score among women correlates .30 with Series 3 of the Rod-and-Frame. Therefore, Hypothesis IV was partially substantiated. Witkin's group found four of eight correlations to be significant when making comparisons similar to those of the present study.

Hypotheses V and VI asserted positive relationships between the Self-concept variables previously described and field-dependence. Inspection of the data shows that neither the Self scores nor the Self-Ideal discrepancy scores are significantly related to any of the measures of field-dependent perceptual performance.

In summary, four of six hypotheses tested in this study were at least partially confirmed, and 26 (29%) of 90 coefficients were found to be statistically significant at the .05 level, a number far greater than chance expectations.

SUMMARY AND CONCLUSIONS

In an attempt to test the general field-dependence hypothesis, six specific hypotheses were examined; four of them represented essentially replications of the Witkin et al. study, while two hypotheses were proposed by the writer as extensions of the original work. Forty-nine male and 48 female summer session university students were given the following perceptual tests: Rod-and-Frame, Jackson's short form of the Embedded Figures and the Chair-Window. Personality tests given were the Self-activity Inventory, the Draw-a-Person, and Holtzman Inkblot. Ten scores derived from these tests were intercorrelated separately for each sex. The results indicated that four of the six hypotheses were at least partially confirmed. Notably, these were the original hypotheses proposed by the Witkin group.

The general field-dependence hypothesis finds considerable support in these data. It is clear that certain aspects of personality are significantly related to perceptual variables. In addition, the patterns of correlations for the sexes reported here show many similarities to the Witkin data. While the present results are not as striking as those in the Witkin study, they nevertheless suggest that there is appreciable validity to the hypothesis, particularly when it is borne in mind that the present study attempted to evaluate the same variables as Witkin but used different instruments in some cases, e.g., the Holtzman Inkblot test and the Chair-Window. The present writer is in agreement with Gruen (1955) who states that the dimension appears not to be as pure, factorially speaking, as Witkin believes. What is needed is the determination of other relevant variables which apparently cut across the dimension. Factorial methods would likely prove helpful in this regard.

Two new findings deserve comment. First, the "depth" at which the personality variables considered in this study were tapped has some relevance in differentiating which class of personality variables are related to perceptual performance. The three measures which presumably tap unconscious levels (DAP, WPC, and MHFA) all show some relation to the field-dependence measures (8 out of 30 correlations significant at .05 level), while none of the correlations between conscious self-evaluations and field-dependence are significant. Second, the Chair-Window test shows striking relationships to other measures considered to be criteria of field-dependence. This test should prove highly useful to investigators in this area of research.

REFERENCES

BARRAT, E. S. Space-visualization factors related to temperament traits. *J. Psychol.*, 1955, **39**, 279–287.

BAUMAN, G. The stability of the individual's mode of perception and perception-personality relationships. Unpublished doctoral dissertation, New York Univer., 1951.

GRUEN, A. The relation of dancing experience and personality to perception. *Psychol. Monogr.*, 1955, **69**, No. 14 (whole No. 399).

HOLTZMAN, W. H. Review of H. A. Witkin, H. B. Lewis, M. Hertzman, K. Machover, P. Meissner, & S. Wapner, *Personality through perception. Amer. J. Psychol.*, 1955, **68**, 501–504.

HOLTZMAN, W. H. Development of an experimental inkblot test, a new departure from the Rorschach. *Amer. Psychologist*, 1956, **11**, 400. (Abstract)

HOLTZMAN, W. H. Objective scoring of projective tests. Paper presented at symposium on objective approaches to personality assessment. February 6–8, 1958, Louisiana State Univer.

JACKSON, D. N. A short form of Witkin's Embedded Figures Test. *J. abnorm. soc. Psychol.*, 1956, **53**, 254–255.

MACHOVER, K. The Figure-Drawing Test. In H. A. Witkin, M. Hertzman, K. Machover, P. B. Meissner, & S. Wapner, *Personality through perception*. New York: Harper, 1954. Pp. 235–254.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception*. New York: Harper, 1954

WORCHEL, P. Adaptability screening of flying personnel. Unpublished manuscript describing research conducted under Air Force Contract No. 18 (600)-913, 1956, Depart. Clinical Psychol., USAF Sch. of Aviat. Med., Randolph Field, Texas.

YOUNG, H. H. Relationships between and reliability estimates of new (Holtzman) inkblot variables and conventional Rorschach scoring categories. *Proc. Okla. Acad. Sci.*, 1958, **38**, 111–115.

Received April 2, 1958.

HOSTILITY AND AVERSION FOR EXPRESSING HOSTILITY IN NEUROTICS AND CONTROLS[1]

H. J. WAHLER

Veterans Administration Hospital, Knoxville, Iowa

FREUD (1930, 1933, 1949) proposed that in neurotics the superego directly or through ego mediation exercises a severe inhibitory effect on socially disapproved impulses, primarily those of an aggressive or sexual nature. In "normal, stable states" he felt that there is a harmonious correspondence between introjected social values mediated by the superego and the ego. The severely repressive action of the superego on disapproved impulses was regarded by Freud as integral to his entire conception of neurosis. Mowrer (1950) challenged this assumption and proposed that instead of the superego's being overly severe in neurotics, it is not severe enough. He felt that while neurotics have learned social values, they are unable or unwilling to apply them in the course of social living. It was his view that many of the failures of psychoanalysis are attributable to therapists' efforts to diminish the severity of the superego when in reality the superego needs to be strengthened.

The need for objective evidence to help resolve such conflicting assumptions and clarify conceptions of the role of social values and inhibitory or aversive tendencies in neurosis seems apparent. However, as Sears (1943) observed, objective studies of psychoanalytic concepts "... may be relatively fruitless so long as those concepts rest in the theoretical framework of psychoanalysis" (p. 143). The superego as a global construct incorporating learned aversion and aversive techniques operating in opposition to certain excitatory tendencies seems to correspond without serious distortion to a conception of approach–avoidance conflict such as that described by Miller (1951).

Freud regarded the superego as developing mainly from childhood experiences in which socially disapproved acts were punished by parents and other authorities. He assumed that cultural values determine the classes of acts which authorities select to punish. With repetition of punishing experiences, he hypothesized that the impulses preceding disapproved acts acquire the property of eliciting what he termed "moral anxiety." He proposed that such anxiety could be reduced or avoided by any of three techniques: conscious avoidance of stimuli evoking the impulses, voluntary suppression, or involuntary repression of impulses. In other terms, goal-oriented action is elicited and punishment occurs with some regularity during the sequence of such actions. With repetition of such acts and punishment, competing aversive tendencies are learned, presumably in association with response-produced stimuli and other elements of the stimulus situation. Both neurotic and nonneurotic people experience such training, but in neurotics, aversive tendencies are more highly overlearned or more highly motivated by anxiety.

As Sears (1943) pointed out, any of the three techniques which Freud felt reduced "moral anxiety," i.e., avoidance, suppression, or repression should "... likewise prevent the occurrence of the activities originally instigated by the impulses." (p. 105). If such aversive tendencies were stronger in neurotic than nonneurotic people, and if impulses instigating the activities were of about equal strength, it should be possible to demonstrate that the overt behavior or performance of neurotics in relevant instances differs from that of nonneurotics. Freud's assumption of greater superego strength in neurotics suggests that they should tend to avoid overt expression of disapproved actions such as certain types of hostile verbalizations to a greater extent than nonneurotic people. Mowrer's assumption of a weaker superego in neurotics suggests the converse: Nonneurotics should tend to avoid overt expression of such disapproved actions to a greater extent than neurotics.

Freud's writings indicate that he regarded *aggressive* impulses to be instinctually deter-

[1] The author wishes to express his gratitude to the following persons for their generous assistance during the completion of this study: I. E. Farber, Arthur L. Benton, M. H. Gordon, Harold Bechtoldt, E. F. Lindquist, S. J. Williamson, and Lynn Roberts.

mined. Apparently, it was his view that the strength of agressive tendencies did not necessarily differ in neurotic and normal people. If it is assumed that hostile tendencies are of equal strength in neurotic and nonneurotic people and that both types of people have acquired aversion for expressing socially disapproved aggression but that such aversive tendencies are stronger in neurotics, then (*a*) neurotics should select fewer socially disapproved hostile alternatives than controls when the situation affords choices between hostile and neutral responses, (*b*) both types of *S*s should select fewer disapproved than nondisapproved hostile statements, and (*c*) neurotics and controls should select nondisapproved hostile *R*s with about the same frequency. On the basis of the hypothesis derived from Mowrer's proposal and the assumption that aggressive tendencies are equally strong in neurotics and controls, neurotics should select *more* disapproved hostile statements than controls.

The above hypotheses do not take into account the fact that aggression is typically defined as harmful action directed against some object. Freud felt that the disapproved aggressive acts most likely to suffer repression were those directed against people. If this were the case, tendencies to avoid the expression of disapproved aggression should be stronger in neurotics than normals when such aggression is directed against people. On the other hand, there does not appear to be any basis for assuming that neurotics are more inclined than nonneurotics to avoid expressing disapproved aggression against nonhuman objects, i.e., animals, insects, plants, etc.

METHOD

Subjects

Twenty-four neurotic[3] *S*s were selected from veterans applying for outpatient treatment at a VA Mental Hygiene Unit. Thirty male veterans were selected as controls from a VA outpatient center who had no neurologic, psychiatric, or psychosomatic diagnoses or complaints. Only *S*s who were less than 50 years of age and who had completed at least six years of school were included in either group. The mean ages of neurotics and controls were 33.3 and 29.1 respectively; the neurotic group was significantly older than controls ($p < .05$). The mean years of education for neurotics was 12.3 and 13.1 for controls. This difference was not significant at the .05 level.

Stimulus Words

The task used was comparable to that described by Watson, Pritzker, and Madison (1955). Sixty-six different sets of scrambled words were used as stimuli. *S*s were required to make one meaningful sentence from each set by dropping one word and rearranging the remaining words. Each set of scrambled words was formed from two sentences; the words in the two sentences were the same except for the verbs. One sentence which could be constructed from each set contained a verb (or verbal adjective) connoting a hostile action or attitude of injury, derogation, or obstruction. The other sentence contained a verb denoting some neutral act such as to see, to observe, to touch, etc. For example, one scrambled set was MAN THE DESTROY TOUCH from which either "Destroy the man" or "Touch the man" could be constructed.

The 132 sentences constituting alternatives in the different scrambled sets of words were selected from a pool of sentences which had been rated on a 9-point scale of social desirability by 19 Air Force reservists. Ebel's (1951) index of reliability for averaged scores was .87 for the ratings of the 132 sentences selected. These sentences were also rated on a 9-point scale of *familiarity* by 43 nurses and nursing assistants. The index of reliability for familiarity ratings was .96.

All of the *neutral* sentences which could be constructed from the scrambled sets of words were selected from a pool of sentences classified as neither approved nor disapproved. The mean social desirability rating of each neutral sentence was greater than 4 but not more than 6. One-third of the hostile sentences were classified as undesirable (mean ratings of 3.9 or less); another third of the hostile sentences were classified as neither approved nor disapproved (mean ratings between 4 and 6); and the remaining third were classified as desirable (mean ratings of 6.1 or more).

Half of the hostile sentences and their corresponding neutral sentences pertained to people. The other half pertained to animals, insects, plants, and inanimate objects. With the sets of words pertaining to human content and the sets pertaining to nonhuman content, one-third of the hostile sentences selected as alternatives by *E* had been independently rated as socially undesirable; another third fell in the classification neither approved nor disapproved; and the remaining third had been rated socially desirable. Thus, there were 11 sets of words out of a total of 66 which pertained to humans and contained hostile alternatives rated undesirable; 11 sets pertained to humans with hostile alternatives rated neither approved nor disapproved; and 11 sets pertained to humans with hostile alternatives rated desirable. Likewise, with *nonhuman* content, there were 11 sets of words with hostile alternatives

[3] Only patients who had neurotic diagnoses in their clinical folders and who were diagnosed as neurotic by the clinical staff were selected for this study. The neurotic group is comprised of the following number of patients in each diagnostic category: 13 anxiety neurosis, 2 psychoneurosis, type unspecified, 1 psychoneurosis, mixed type, 2 neurasthenia, 3 anxiety hysteria, 1 nervous stomach, 1 duodenal ulcer, and 1 conversion reaction.

falling in each of the three social desirability classifications. The following are examples of groups of words which fell in the six classifications (not in scrambled order): *Human content*: touch, destroy the enemy; slap, see his back; dislike, note the neighbors. *Nonhuman content*: see, slap the flies; watch, punch the clock; and break, observe the windows.

The mean familiarity rating of each *hostile* sentence from a scrambled group of words was subtracted from the mean familiarity rating of the corresponding *neutral* sentence. Differences among the means of these values for items falling in each of the six classifications were evaluated statistically by a two-factor analysis of variance. None of the *F*s for main effects or interaction was significant, indicating that there were no significant differences in the average relative familiarity of alternatives in any of the six classifications.

Two different permutations of the words in each scrambled group were formed. This was done to counterbalance the relative positions of the hostile and nonhostile verbs within each group of words. These two sets of word permutations were also arranged in three presentation orders. Equal numbers of *S*s from neurotic and control groups were assigned to the six permutation-by-order classifications in their order of acceptance for the study.

Apparatus and Procedure

The scrambled groups of words were typed on 5 x 8 cards in capital letters. The grammatical object of the two sentences which could be formed from each scrambled group was underlined in black ink. *S*s were instructed never to omit the underlined word. This eliminated the possibility of *S*s' forming sentences other than the two choices intended.

The presentation apparatus consisted of a two-foot square board mounted on a base with a small window cut in the lower center. A sliding door could be raised or lowered behind the window. The cards containing the scrambled groups were placed in a rack directly behind the window and were presented one at a time as fast as *S*s could respond and *E* could operate the apparatus. The average response latency for all *S*s was 4.28 seconds, with a range from 1.1 to 96.8 seconds; *E* required approximately 5 sec. to record responses and operate the apparatus.

Individual *S*s were brought into *E*'s office and seated before the apparatus. Clinical *S*s were told that the procedure was part of their intake examination. Controls were told that the hospital wanted brief personality evaluations of all outpatients and were requested to take some "tests" for that purpose; they were not told that they were participating in a research project. All *S*s were told that they would be shown a few words which were mixed up and that they were to tell *E* as quickly as possible the first short sentence which came to mind when they left out one of the words shown which was not underlined. They were told the task was to determine how quickly they could respond. After instructions, the stimulus items were presented individually without interruption in the prearranged order assigned to each *S*.

TABLE 1

MEAN FREQUENCY OF HOSTILE RESPONSE SELECTIONS OF NEUROTIC AND CONTROL GROUPS FOR SOCIAL DESIRABILITY AND CONTENT CONDITIONS

	Neurotics		Controls	
	Human	Non-human	Human	Non-human
Socially desirable	6.54	6.54	5.73	6.37
Neither approved nor disapproved	7.62	6.17	6.83	6.33
Socially undesirable	5.12	5.08	4.33	5.37

RESULTS

The scores used in the analyses consisted of the frequency with which hostile statements were selected in each of the six content-by-social desirability categories. Since there were 11 scrambled groups of words in each category, the maximum score in each category was 11. A score of 11 meant an *S* selected only hostile statements in a particular category. A score of zero meant an *S* selected only neutral statements in a particular category.

The mean frequency with which neurotics and controls selected hostile responses in each of the six categories is given in Table 1. These data were evaluated statistically by means of analysis of variance (Lindquist, 1953, Type VI). The summary of this analysis is given in Table 2.

Inspection of Table 2 shows that the triple interaction was not significant. The nonsignificant two-factor social desirability-by-groups interaction suggests that any effects attributable to the different social desirability values of the hostile alternatives may be regarded as parallel (except for chance deviations) for the neurotic and control groups with both human and nonhuman content. These parallel trends may be observed from the means in Table 1. With human content, neurotics selected more hostile responses than normals in all three social desirability classifications. With nonhuman content, the mean frequencies of hostile selections were nearly identical for neurotics and controls in all three social desirability classifications.

Table 2 shows that the main effect of social desirability was highly significant ($p < .001$). Since this factor interacted significantly with content, the differences among the three means

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE

Source	*df*	*SS*	*Ms*	*F*	*P*
Between *Ss*	53	288.09			
Groups	1	9.96	9.96	1.86	—
error (b)	52	278.13	5.35		
Within *Ss*	270	914.83			
Social desirability	2	180.19	90.10	25.10	.001
Content	1	0.003	0.003	—	—
Social desirability by content	2	34.83	17.42	9.22	.001
Content by groups	1	15.80	15.80	7.31	.01
Social desirability by groups	2	.81	.40		—
Social desirability by content by groups	2	.70	.35	—	—
error (w)	260	682.50			
$error_1$ (w)	104	373.33	3.59		
$error_2$ (w)	52	112.36	2.16		
$error_3$ (w)	104	196.80	1.89		
Total	323	1202.95			

TABLE 3
MEAN NUMBER OF HOSTILE RESPONSES SELECTED BY NEUROTICS AND CONTROLS WITH HUMAN AND NONHUMAN CONTENT

	Neurotics	Controls	t^a for Horizontal Differences	*P*
Human content	6.43	5.63	2.28	.05
Nonhuman content	5.93	6.02	.37	—
t^a for vertical differences	2.16	1.68		
P	.05	.10		

[a] 52 *df*.

(averaged across content and groups) lack generality. Tests of the significance of the differences between means in each of the three social desirability classifications with human content showed that both neurotics and controls selected significantly fewer disapproved hostile statements than hostile statements classified as desirable or neither approved nor disapproved. The same results were obtained with nonhuman content. These differences were all significant at the .01 (or less) level of significance.

Another significant effect shown in Table 2 is the content-by-groups interaction ($p < .01$). Table 3 contains the mean frequency with which hostile statements were selected by neurotic and control *Ss*, averaged across the three social desirability classifications for both human and nonhuman content. It may be seen from Table 3 that, on the average, neurotics selected significantly more hostile statements directed against humans than did controls. With hostility directed against nonhuman objects, the average group difference does not approach significance.

DISCUSSION

The findings may be interpreted relatively simply by means of three definitional assumptions: (*a*) frequency of hostile selections is a monotonic function of the strength of differential hostile tendencies in individuals. (*b*) Aversion for expressing hostility is evoked when hostile responses are socially disapproved and not evoked when hostile responses are not disapproved; the stronger the aversive tendencies, the greater the frequency of neutral selections. (*c*) Hostile and aversive tendencies activated simultaneously summate algebraically.

With *nonhuman* content, both groups selected significantly fewer disapproved than nondisapproved hostile responses. The differences between these frequencies (number of nondisapproved minus number of disapproved selections) were approximately the same for both groups. Likewise, neurotic and control *Ss* did not differ significantly in terms of the mean frequency (or level) of disapproved and nondisapproved hostile selections. Thus, there is no evidence that hostility when associated with nonhuman content is stronger in neurotic than control *Ss* or that aversion for expressing disapproved hostility is on the average of different strength in the two groups.

When content pertained to *humans*, neurotic *Ss* selected significantly more disapproved and nondisapproved hostile *R*s than controls. The differences between the mean number of disapproved and nondisapproved hostile selections made by neurotic and controls *Ss* were approximately equal (frequency of nondisapproved minus disapproved). These findings suggest that when hostility is directed against humans, hostile tendencies are stronger in neurotics than in controls but that aversion for expressing disapproved hostility is of approximately equal strength in both groups.

If the suppressive action of the superego is related to overt behavior, as Sears (1943) suggested, the findings do not support Freud's proposition that such suppressive tendencies

are stronger in neurotics in relation to the expression of disapproved hostility. The findings also fail to support Mowrer's proposal that the superego is weaker in neurotic people in relation to the expression of disapproved aggression. While neurotics selected more disapproved hostile *R*s than controls with human content, they also selected relatively more nondisapproved hostile *R*s. On the assumption from conflict theory that approach and avoidance tendencies summate albegraically, the more frequent selection of disapproved *R*s by neurotics may be more readily interpreted as a function of stronger hostile rather than weaker aversive tendencies. As is well known, the interpretation that hostility is stronger in neurotics than normals is consistent with considerable clinical observation reported by Horney (1938), Adler (1927), Sullivan (1947), Cameron (1947), and others. It is also noteworthy that all of these authors tend to emphasize the view that neurotic hostility is typically directed against people.

Although the results failed to provide positive support for Freud's assumption of relatively stronger aversive tendencies in neurotics, it must be stressed that the generality of this conclusion is limited to a subpopulation of neurotics in general and to only one type of disapproved behavior. The possibility remains that neurotics may show stronger aversion than controls for expressing other types of disapproved behavior. For example, Fenichel (1945) stated that in neurotics one type of aversive mechanism, repression, may occur more typically in conjunction with sexual than with hostile impulses. Fenichel also stated that certain types of neurotics such as hysterics are more prone to exhibit repressive tendencies than such other types as anxiety neurotics, who composed the majority of the experimental group in this study.

SUMMARY

The purpose of this study was to investigate, by means of objective methodology, certain aspects of the conflicting assumptions of Freud and Mowrer regarding the strength of the superego in neurotics. Freud proposed that the superego in neurotics is overly severe in its suppressive effect on socially disapproved impulses. Mowrer contended that in the case of neurotics the superego is not severe enough.

To investigate these proposals, neurotic and control *S*s were required to construct and report one sentence each from a series of scrambled words. The scrambled words were chosen so that two alternative sentences could be formed, one being hostile and the other neutral. Some of the hostile alternatives had been independently rated as socially undesirable and some as desirable or as neither socially approved nor disapproved. All neutral alternatives used were rated neither approved nor disapproved. Half of the series of scrambled words pertained to human content and half to nonhuman content. Relative familiarity of items in each of these categories did not differ significantly.

On the assumption, implicit in Freud's writings on aggression, that hostile tendencies are no stronger in neurotic than nonneurotic people, it was hypothesized that if the superego were stronger in neurotics, they should select fewer socially disapproved hostile alternatives than controls and that the two groups should not differ in terms of the frequency of nondisapproved hostile selections. Since hostility directed against people is assumed to be most likely to suffer repression in neurotics, it was also hypothesized that in comparison with controls, neurotics should select relatively fewer disapproved hostile statements directed against people than against nonhuman objects.

On the basis of Mowrer's proposal, it was hypothesized that neurotics should select more socially disapproved hostile alternatives than should controls, regardless of content.

The findings showed that both neurotic and control groups selected significantly fewer socially disapproved than nondisapproved statements. When responses of the two groups were compared, it was found that neurotics and controls selected approximately the same number of socially disapproved hostile statements and about the same number of nondisapproved hostile statements when content pertained to nonhuman objects. However, when content pertained to humans, neurotics selected significantly more disapproved and nondisapproved hostile statements than did controls.

These results were interpreted in terms of an assumption of conflict theory that conflicting response tendencies summate algebrai-

cally. On this assumption, the findings suggest that (*a*) aversion for expressing disapproved hostility is, on the average, approximately equally strong in members of neurotic and control groups; (*b*) on the average, hostile tendencies are stronger in neurotics than in controls and (*c*) stronger hostility in neurotics is specifically associated with human content.

Limitations in the generality of the findings and interpretations were discussed in terms of the type of behavior studied as well as the sample of neurotics.

REFERENCES

ADLER, A. *Understanding human nature.* New York: Garden City, 1927.

CAMERON, N. A. *The psychology of behavior disorders.* New York: Houghton Mifflin, 1947.

EBEL, R. L. Estimation of the reliability of ratings. *Psychometrika,* 1951, **16,** 407–424.

FENICHEL, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.

FREUD, S. *Civilization and its discontents.* London: Hogarth Press, 1930.

FREUD, S. *New introductory lectures on psychoanalysis.* New York: Norton, 1933.

FREUD, S. *An outline of psychoanalysis.* New York: Norton, 1949.

HORNEY, KAREN. *The neurotic personality of our time.* New York: Norton, 1937.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.

MILLER, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. *J. Pers.,* 1951, **20,** 82–100.

MOWRER, O. H. *Learning theory and personality dynamics.* New York: Ronald, 1950.

SEARS, R. R. *Survey of objective studies of psychoanalytic concepts.* New York: Soc. Sci. Res. Council, 1943.

SULLIVAN, H. S. *Conceptions of modern psychiatry.* Washington, D. C.: William Alanson White Psychiatric Found., 1947.

WATSON, R. E., PRITZKER, L., & MADISON, P. Hostility in neurotics and normals. *J. abnorm. soc. Psychol.,* 1955, **50,** 36–41.

Received June 6, 1958.

PERSONALITY FACTORS IN MOTOR LEARNING AND REMINISCENCE[1, 2]

OAKLEY S. RAY[3]
University of Pittsburgh

EYSENCK (1956) has employed a rotary pursuit paradigm to test the hypothesis that "extraverts should show a higher degree of reminiscence than intraverts" (p. 328), which he feels follows from a postulate involving these personality factors. According to Eysenck (1955):

> Individuals in whom reactive inhibition is generated quickly, in whom strong reactive inhibitions are generated, and in whom reactive inhibition is dissipated slowly are thereby predisposed to develop extraverted patterns of behavior and to develop hysterico-psychopathic disorders in cases of neurotic breakdown; conversely, individuals in whom reactive inhibition has developed slowly, in whom weak reactive inhibitions are generated, and in whom reactive inhibition is dissipated quickly, are thereby predisposed to develop introverted patterns of behavior and to develop dysthymic disorders in cases of neurotic breakdown (p. 35).

To test this prediction Eysenck used a single rest interval and correlated the extroversion scores of 50 subjects (*S*s) with their performance increases over this rest interval. He obtained a significant correlation, +.29, thus upholding his prediction. The present study was designed to test the generality of Eysenck's prediction by using several rest intervals and to investigate the differential rate of I_R accumulation over the course of original learning.

Although Eysenck (1956) states that "the argument [that extroverts will show more reminiscence than introverts] is probably too obvious to require much elaboration" (p. 328), this hypothesis appears to be less than a direct inference from his postulate.

Eysenck (1956) argues that "massed practice produces inhibition; inhibition is generated more quickly in extraverts. . . . Reminiscence is produced by the dissipation of the accumulated inhibition during a rest period. . . . Extraverts who have accumulated more inhibition will therefore have more inhibition to dissipate and will consequently show greater reminiscence effects" (p. 328). It must be noted however, that Eysenck also states that introverts dissipate I_R more rapidly than extroverts. Over a given period of time, then, more inhibition may be dissipated by the rapid dissipation of I_R in introverts than by its slow dissipation in extroverts—even though the extroverts had more I_R to dissipate at the beginning of the period. It is thus possible for introverts to show more reminiscence than extroverts, according to Eysenck's formulation. Eysenck's postulate leads to the direct prediction that when both groups have completely dissipated their I_R (as after long rest intervals), the extroverts will show greater performance increases (reminiscence), since they originally had more I_R. His theory would be compatible, however, with divergent results for short rest intervals.

Although Eysenck makes no statements regarding the performance curves for introverts and extroverts during the original work period, it is possible to predict directly from his postulate that introverts will have a higher performance level than extroverts, since extroverts build up I_R (a response depressor) more rapidly than introverts.

METHOD

Apparatus

The pursuit rotor used was a modified Lafayette model with a Prestwood turntable 10 in. in diameter that rotated clockwise at 60 rpm. The target was a silver disc ¾ in. in diameter flush with the turntable surface. The target's inside edge was 2⅞ in. from the center of rotation. The stylus handle was 3¾ in. long and ⅝ in. in diameter. The stylus arm was 6⅝ in. in length and the last 1⅛ in. had been bent at a 90-degree angle to the rest of the stylus. The stylus tip was convex in shape, and when resting free it exerted a pressure of ½ oz.

[1] Submitted to the graduate faculty in the Division of the Natural Sciences at the University of Pittsburgh in partial fulfillment of the requirements for the degree of doctor of philosophy. The author wishes to express his appreciation to G. Wischner, chairman, and to A. W. Bendig, H. W. Braun, A. D. Lazovik, and J. Matthews, the other members of his committee.

[2] Portions of this study were presented at the APA meetings in Washington, D. C., August 1958.

[3] The author is presently a National Institute of Mental Health Postdoctorate Fellow of the USPHS at the VA Research Laboratories in Neuropsychiatry, Leech Farm VAH, Pittsburgh.

Two decade interval timers and two standard electric clocks calibrated to .01 sec. were used in recording the *S*'s performance. The decade interval timers operated on a 20-sec. cycle so that one clock recorded the number of seconds an *S* stayed on target during a 20-sec. interval, and then the second clock recorded for a 20-sec. period. The only manual operations required were the recording of the number of seconds on target for each 20-sec. interval, and resetting the clock to zero before it began recording again. Scores were recorded to .01 sec.

The recording apparatus and experimenter were screened so that the experimenter could watch the *S* but the *S* could see neither the experimenter nor the apparatus. Illumination was artificial and constant for all *S*s.

Procedure

All *S*s were given standard instructions prior to 5-min. continuous work on the pursuit rotor.[4] A continuous work period is essential, since the interposition of brief rest intervals during the work period would allow some I_R to dissipate. This point is critical in a test of Eysenck's postulate because two hypotheses are involved: extroverts build up I_R more rapidly than introverts, and they dissipate it more slowly than introverts. If at the end of a discontinuous work period the performance levels are measured and the introverts have done better than the extroverts, it would be impossible to say whether the difference had been caused by the more rapid growth of I_R during work in the extrovert group, or by its slower dissipation in the extrovert group during the rest intervals. Following this 5-min. work period, *S*s were given 0, 30, 60, 180, 600, or 900 sec. rest prior to a 2-min. continuous work period.

Subjects

The *S*s were 240 students enrolled in undergraduate psychology courses drawn from a total population of 515 students who took the Maudsley Personality Inventory (MPI) (Bendig, 1957; Eysenck, 1956). Scores of 25 and below, the bottom 30% of the extroversion scores, placed the *S*s in the introvert group. Scores of 35 and above placed the *S*s in the extrovert group, the top 30% of the extroversion scores. All *S*s in these groups whose age was 30 or more were eliminated.

Correlations between the two scales of the MPI, extroversion and neuroticism, were obtained from the population separately for males and females. Ninety-one females had a correlation of −.19 between extroversion and neuroticism ($p < .10$), while 126 males had a correlation of −.24 ($p < .05$) on these measures.

Since there was a correlation of −.22 between extroversion and neuroticism, the median neuroticism score in the introvert group was 23, while in the extrovert group the median neuroticism score was 16. The difference between the neuroticism scores for these groups was significant ($p < .01$) using chi square.

Each *S* within the major experimental groups, introvert and extrovert, was placed in a high or low neuroticism group by dividing each of the major groups at their own median score on neuroticism. It should be emphasized that the *S*s were selected solely on the basis of their extreme extroversion scores. After they were categorized into an introvert and an extrovert group, they were then subcategorized on the basis of their neuroticism scores.

There were 60 *S*s in each of these four groups: introvert, high neuroticism; introvert, low neuroticism; extrovert, high neuroticism; extrovert, low neuroticism.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF EXTROVERSION AND NEUROTICISM SCORES FOR THE 24 EXPERIMENTAL GROUPS

Rest Condition	0	30	60	180	600	900
Extrovert, High Neuroticism Extroversion:						
M	38.10	37.50	37.40	38.40	38.80	37.80
σ	2.38	2.91	2.11	3.54	3.28	2.34
Neuroticism:						
M	26.70	24.60	22.00	24.20	28.00	26.80
σ	8.69	6.99	4.14	4.91	7.71	5.35
Extrovert, Low Neuroticism Extroversion:						
M	38.50	37.90	38.80	39.20	38.70	37.50
σ	2.15	2.88	3.81	2.81	2.76	2.50
Neuroticism:						
M	9.60	9.70	7.60	10.30	8.90	10.00
σ	5.31	4.52	4.61	4.26	4.28	4.05
Introvert, High Neuroticism Extroversion:						
M	16.60	18.80	18.80	18.10	16.40	18.00
σ	5.98	4.21	3.65	4.22	5.82	6.39
Neuroticism:						
M	33.70	32.60	31.00	30.00	35.60	32.40
σ	4.15	3.16	3.43	5.05	4.31	7.59
Introvert, Low Neuroticism Extroversion:						
M	20.20	20.20	19.50	21.00	20.00	19.10
σ	2.64	4.75	5.00	3.55	4.85	4.26
Neuroticism:						
M	15.00	14.00	12.90	12.80	15.40	14.20
σ	5.96	4.71	7.16	6.37	5.02	5.42

[4] *Instructions to Subject*: Your task is to keep the tip of the stylus on the silver disk, you do it like this [demonstrate]. Just move your arm right around with it. Don't lift the stylus from the turntable; just try to stay on top of the disk. The stylus is hinged so that it rests lightly on the turntable [illustrate]. Don't press down on the stylus or hold it [illustrate]. Keep your hand on the black handle only. Don't touch the metal stylus. Are there any questions? [Hand stylus to subject.] Is that the hand you write with? O.K., just stand here [in front of the turntable], hold the stylus over the turntable, don't move your hand. Wait for my signal—don't start until I say "Go." Don't lift the stylus from the turntable until I say "Stop."

Ten *Ss* from each group, 7 male and 3 female, were assigned to each of the six experimental conditions.

The means and standard deviations of the extroversion and neuroticism scores for the 24 subgroups are contained in Table 1. Introvert and extrovert groups (120 *Ss* each) differed significantly in both extroversion and neuroticism scores.

RESULTS

Original learning. Original learning for all 240 *Ss* consisted of 5-min. continuous performance on the pursuit rotor. Original learning data were analyzed in terms of time on target per minute, as shown in Fig. 1. Figure 1 presents the original performance curves of each of the four experimental groups ($N = 60$, each group). At all points the introvert, low neuroticism group had superior performance and the extrovert, low neuroticism group the poorest performance.

Table 2 shows that the difference in performance level between introverts and extroverts approaches but does not reach significance, and that the difference in slopes is significant at the .01 level. It should be noted here that although the females were consistently poorer than the males, the ordering of their groups from best performance to worst was the same as for the male groups: introvert, low neuroticism; introvert, high neuroticism; extrovert, high neuroticism; extrovert, low neuroticism.

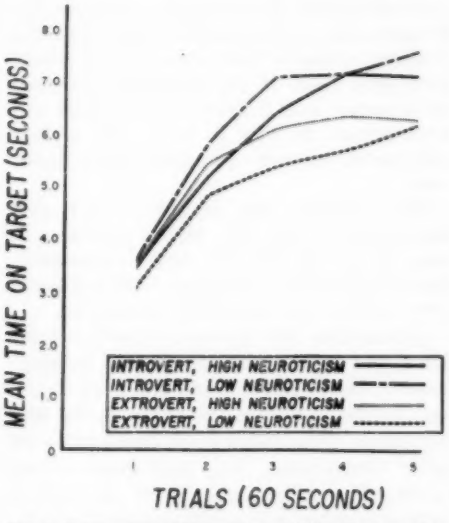

FIG. 1. PERFORMANCE CURVES FOR 5 MIN. CONTINUOUS PRACTICE FOR 60 INTROVERTS WITH HIGH AND 60 INTROVERTS WITH LOW NEUROTICISM SCORES, AND 60 EXTROVERTS WITH HIGH AND 60 EXTROVERTS WITH LOW NEUROTICISM SCORES

TABLE 2

REPEATED MEASURES ANALYSIS OF VARIANCE OF ORIGINAL PERFORMANCE SCORES CLASSIFIED BY THE VARIABLES EXTROVERSION, NEUROTICISM, AND REST CONDITION

Sources of Variation	*df*	Mean Square	*F*
Between *Ss*	239	97.16	
Groups	23	34.48	
Extroversion	1	182.70	1.78
Neuroticism	1	1.12	
Rest Condition	5	30.45	
Extroversion × Neuroticism	1	57.27	
Extroversion × Rest Condition	5	9.77	
Neuroticism × Rest Condition	5	6.75	
Extroversion × Neuroticism × Rest Condition	5	63.42	
Between *Ss* Within Groups	216	102.49	
Within *Ss*	960	6.59	4.28
Trials	4	471.79	306.36**
Trials × Extroversion	4	12.04	7.82**
Trials × Neuroticism	4	2.43	1.58
Trials × Rest Condition	20	4.26	2.77**
Trials × Extroversion × Rest Condition	20	12.97	8.42**
Trials × Neuroticism × Rest Condition	20	4.61	2.99*
Trials × Extroversion × Neuroticism	4	535.44	347.69**
Trials × Extroversion × Neuroticism × Rest Condition	20	23.80	15.45**
Pooled *Ss* × Trials	864	1.54	

* $p < .05$.
** $p < .01$.

Reminiscence. The reminiscence scores used hereafter refer to the absolute difference between the mean number of seconds on target during the last 20-sec. prerest work period and the mean time on target during the first 20-sec. postrest work period for the same group.

Figure 2 presents the reminiscence results for the four experimental groups. Each point is based on the mean score of 10 *Ss* in each of the six rest intervals. It should be noted that all four control groups showed a drop in performance level from the "prerest" to the "postrest" trial. No trends in amount of reminiscence

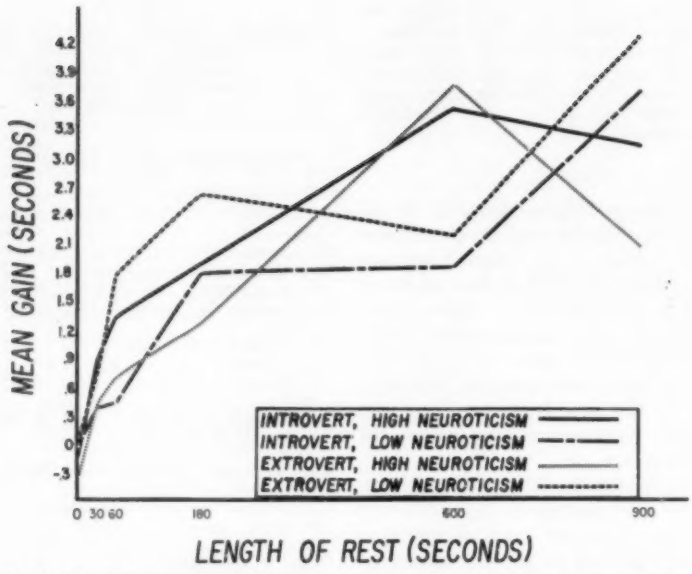

FIG. 2. MEAN PERFORMANCE INCREASE AFTER REST FOR INTROVERTS WITH HIGH AND LOW NEUROTICISM SCORES, AND FOR EXTROVERTS WITH HIGH AND LOW NEUROTICISM SCORES; TEN *Ss* IN EACH GROUP AT EACH POINT

seem apparent from the figure and an analysis of covariance indicated that there were no significant differences in reminiscence over the five rest intervals for the four experimental groups.

DISCUSSION

Original learning. The prediction that introverts would have a higher performance level in a continuous work task than extroverts was not upheld, although the difference in slopes of the performance curves of these two groups was significant as Fig. 1 and Table 2 indicate. It would seem that these results support Eysenck's hypothesis that I_R builds up more rapidly in extroverts than in introverts. No relationship between neuroticism score and performance is indicated by either Fig. 1 or Table 2.

Reminiscence. It was predicted that the absolute amount of performance gain when maximum reminiscence had occurred would be greater in extroverts than in introverts, since the extroverts had more I_R to dissipate. Neither group apparently reached asymptotic reminiscence levels following the longest rest intervals used in this study, so it is not possible to handle this prediction definitively. From the evidence available here, however, it would seem that there is no relationship between reminiscence over varying rest intervals and extroversion. Neither is there any evident relationship between neuroticism and reminiscence.

After 600-sec. rest, Eysenck found that neuroticism scores were better predictors of amount of reminiscence than extroversion scores, neuroticism correlating +.40 and extroversion only +.29 with performance gain.

Correlations were obtained between neuroticism and extroversion scores and reminiscence after 600-sec. rest and also after 30-sec. rest for the *Ss* used in this study. In the 600-sec. group, extroversion scores correlated +.08 with reminiscence, and neuroticism scores +.09 with reminiscence ($N = 40$). While neither correlation is significant, both correlations are in the direction reported by Eysenck. In the 30-sec. group, neuroticism correlated −.03 and extroversion −.01 with reminiscence ($N = 40$).

SUMMARY

One hundred twenty introverts and 120 extroverts (selected on the basis of extreme scores on the Maudsley Personality Inventory) were given 5-min. continuous work on a pursuit rotor prior to rest periods of 0, 30, 60, 180, 600, or 900 sec. Two-minutes postrest work was then given.

Over the original learning period introverts had a significantly higher performance trend than extroverts. There were no significant differences in reminiscence between the introvert and extrovert groups. No relationships were found between neuroticism scores and either original learning or reminiscence.

The superior performance trend of the introverts during original learning was interpreted as supporting Eysenck's hypothesis that extroverts build up I_R more rapidly than introverts. Failure of the extrovert groups to show significantly more reminiscence than the introvert groups suggests that either Eysenck's formulation regarding the rate of dissipation of I_R in introverts and extroverts is incorrect, or that the technique used here is not a good way to test it.

REFERENCES

BENDIG, A. W. College norms for the Pittsburgh revision of the Maudsley Personality Inventory. *Psychol. Newsltr., NYU*, in press.

EYSENCK, H. J. A dynamic theory of anxiety and hysteria. *J. ment. Sci.*, 1955, **101**, 28–51.

EYSENCK, H. J. Reminiscence, drive, and personality theory. *J. abnorm. soc. Psychol.*, 1956, **53**, 328–333.

Received June 18, 1958.

EFFECTS OF CONFIDENCE AND MOTIVATION ON SUSCEPTIBILITY TO INFORMATIONAL SOCIAL INFLUENCE[1]

FRANCIS J. DI VESTA

Syracuse University

RESEARCH on conformity behavior suggests that the influence of social pressure may be derived from at least two general sources. One of these, a group structure variable, has been described by Thibaut and Strickland (1956) as a *group set* and by Deutsch and Gerard (1955) as *normative social influence*. Under conditions that maximize this variable, individuals make judgments while functioning as bona fide members of a group. An individual's compliance, when his judgments disagree with those of the majority, is assumed to provide a means of maintaining membership in the group, of contributing to its code and norms, and of enhancing its cohesive qualities.

The second pressure influence, a task variable, is described respectively by these same investigators as a *task set* or *informational social influence*. Under conditions that maximize this variable, group membership is relatively unimportant, and group influences are only subtly or indirectly created, if at all. Compliance occurs as a result of the individual's need for attaining cognitive clarity of his environment. He resorts to the judgments of the other members (to use Thibaut and Strickland's phrase) "as mediators of fact."

The concern in the present study is with the task variable. The general hypothesis holds that under conditions of informational social influence, the individual's acceptance of the judgments of others varies as a function of (*a*) the relative confidence he has in his own judgment as compared with the judgments of others; (*b*) his motivational orientation toward the task with which he is confronted.

[1] This is one of a series of reports of a program of research on conformity conducted in the Department of Psychology at Syracuse University. This research was supported in whole or in part by the United States Air Force under Contract AF 18(603)-20 monitored by the Air Force Office of Scientific Research of the Air Research and Development Command. Permission is granted for reproduction, translation, publication, use and disposal of this article in whole or in part by or for the United States Government.

METHOD

Subjects

A total of 385 *Ss* were used in the experiments described here. An additional 105 *Ss* were used in a replication of one of the experiments. All *Ss* were sophomores, enrolled in the introductory psychology course at Syracuse University, who participated as part of the course requirement. The study was defined to the *Ss* as an experimental study of perception.

Apparatus

Conformity was measured through a modified Crutchfield apparatus that accommodated five *Ss*. Briefly, it consisted of five adjacent panels on each of which were five rows of one red and four green signal lights and a row of four mercury switches. The five rows of lights were labeled A, B, C, D, and E, to represent each of the members in the group. One of the lights in the column of red lights was used to signal *S* when he was to respond for a particular task. The four columns of green lights represented the item alternatives; the one that lit in any row supposedly informed *S* of the response of one of the other *Ss*. The panels were separated by plywood separates to form a three-sided enclosure. The switches on *Ss*' panels were wired separately to the *E*'s control panel, located behind the *Ss*. This wiring scheme permitted *E* to receive information separately from each *S*. The lights on *Ss*' panels were wired in parallel to permit *E* to send identical information to all five *Ss* simultaneously. The signals were identical for all *Ss* on any one task. Detailed description of the apparatus is available elsewhere (Crutchfield, 1955; Di Vesta, 1958; Nakamura, 1958).

Measure of Conformity

The scale used for measuring conformity consisted of 86 items mounted on individual slides and projected to a screen in front of the apparatus by a 35-mm. projector. Thirty-six of the items were critical items. Of these, 8 required attitudinal judgments, for example, "I try to avoid criticism whenever possible"; 12 required informational judgments, for example, "How many stripes does a Master Sergeant wear"; and 16 required perceptual judgments, for example, "Which of the alternative lines is the same length as the standard line?" There were four alternatives to each item. *Ss* always responded in the last position for the critical items and the erroneous signals for these items were always discrepant by one or two positions from the veridical or modal response. There were 16 bogus items of the same types as in the critical items. On these, incorrect signals were presented to the *S* when he was responding from one of the other four positions. There were 34 filler items in which the signals coincided with

the veridical or modal response. *S* might be in any one of the five positions when responding to these items. On the filler and bogus items the signals were occasionally varied. These procedures were used to lend credibility to the instructions that the signals represented the responses of the other members of the group.

The scale was divided into two parallel forms referred to as Form A and Form B, constructed to be equivalent with respect to total length, number of critical items, item type, and placement of items in the series. The forms were administered consecutively as a single scale without interruption. The scale was modified for use with the pretraining variations as described below.

A response to a critical item was scored as reflecting conforming behavior if it deviated from the veridical or modal response and coincided with the supposed responses of others as presented to *S* in the panel of lights. The maximum possible score was 36, 18 on each form of the measure.

Procedure

The design involved four basic variations in procedure. (*a*) Two sets of groups were used to study the effects of certain *pretraining variations*. The first was used to study the persistence of conformity behavior by comparing performance, under social influence conditions, on the first half of the critical trials with that on the second half. With the second, the number of critical trials on the first half of the test was varied to investigate the effects of pretraining on subsequent conformity behavior. (*b*) An additional condition was introduced to study the effects of *personal involvement*. Instructions were manipulated so that one group received the basic instructions for the social influence situation, and a second group received the added information that the test was highly correlated with intelligence. (*c*) The effects of high and low *evaluation* of performance on subsequent conformity behavior were studied in an appropriate experimental variation. (*d*) Finally, appropriate *control* conditions were established to isolate the effects of experimental variables.

Pretraining variations. This phase of experimentation was designed to study the effects of first experiences (pretraining) in the social influence situation on subsequent conformity behavior. In one set of conditions designed to study the persistence of the social influence effects, *Ss* ($N = 80$) were required to make judgments of the tasks on Form A and Form B of the test under influence conditions. The presentations were counterbalanced so that half the groups responded to Form A followed by Form B, and the alternate half responded to Form B followed by Form A. This procedure was replicated with 105 additional *Ss*.

A second set of conditions was designed to study the effects of different degrees of majority error during the initial experiences in the social influence situation on subsequent conformity behavior. Each of five experimental groups ($N = 27$ to 29 in each) was presented one of five series of Form A modified as follows to provide differential pretraining: All items and positions of the *S* when responding to a specific item were retained as in the original form. In the first series all majority reports (inputs) were made to coincide with the veridical or modal response. The remaining four series were respectively developed by converting 6, 12, 18, and 24 items to critical items. The six items selected for the second series were retained for the last three series; those in the third series were retained for the fourth and fifth series; and so on. Form B, the criterion measure, was retained exactly as in the original form. It was administered immediately following the modified Form A series without interruption or further instruction.

The personal involvement variation. This procedure was designed to study the effect of sets, induced by instructions, on conformity. The experimental group ($N = 30$) in this variation differed from the basic social influence condition only with respect to the instructions given prior to the administration of the task. *Ss* were told that scores on the tasks they were about to perform were correlated with intelligence and that the correlation was a high one. The responses of this group were compared with those of a control group ($N = 25$) responding without the instruction (see C-3 under *Controls*). In this and following variation, together with the control group C-3, *Ss* responded to the tasks from memory after stimuli were exposed 20 sec.

The evaluation variation. This was a follow-up of the personal involvement variation and was designed to investigate the effects of evaluation of performance on subsequent conformity behavior. *Ss* ($N = 50$) were instructed at the outset that performance on the tasks was a reflection of intelligence, as they were in the personal involvement variation. In addition, they were told that midway through the series of trials they would be informed of their scores. Immediately after the first series (Form A), four minutes were taken to "score" the responses of each *S* and to give the *S* his relative standing in his group. In reality, the responses were not scored, and *Ss* received only fictitious reports of their performance. Three *Ss* in one group received a "highest" or "among the highest" evaluation of their performance, and two *Ss* received a "lowest" or "among the lowest" evaluation. These grades were assigned at random. In alternate groups, the number of "high" and "low" evaluations within the group were reversed so that half the *Ss* ($N = 25$) received one of the evaluations, and the second half ($N = 25$) received the other. Following these evaluations, *E* proceeded with the second series of trials.

Controls. Three control groups were used. In the first (C-1, $N = 27$), *Ss* responded by means of switches to the series of tasks under "private" conditions, i.e., without signals. This control provided a measure of the extent to which errors occurred on the critical items in the absence of information concerning others' responses.

The second control group (C-2, $N = 23$), responded to the series of trials with the same schedule of lights as used in the experimental groups. Unlike the experimental groups, however, *Ss* in this condition were informed that the signal (green light) in each row of lights was an indication that the *S* in the booth corresponding to that row had responded; that the lights were *controlled by E* and were merely *E*'s way of letting *S* know when it was his turn to respond; that the lights had

nothing to do with the judgment of any *S*; and that a specific light position was *selected by chance*.

The third control group (C-3, $N = 25$) responded to the measure with the same schedule of lights as in C-2. However, the signals were attributed to the *S* in the booth corresponding to the row in which the signal appeared. Thus, the *S* was led to believe the signals represented the responses of the others. This condition was used as the control for the personal involvement variation.

RESULTS

The analysis in each case concerns the number of errors made by *S*s which coincided with the background reports represented by *E*'s schedule of signals.

The Control Situations

The mean number of errors for the C-1 group (signals absent) was 3.67; that is, *S*s who are not subjected to influence pressure of any kind make approximately 10% of the maximum possible error score. The range of errors was from 1 to 7. The mean number of errors for the C-2 group (signal present but not attributed to group members) was 5.87. The range of errors was 2 to 29, reflecting the effect of a combination of chance errors made in perceptual judgment (see C-1) plus compliance to an arbitrary sequence of lights. Since the distribution for the C-2 control group was markedly skewed, the nonparametric Mann-Whitney *U* test was employed to evaluate the significance of the difference. The obtained value of *z* was 4.05, which has a probability under H_0 of $< .001$. The decision is to reject the null hypothesis and to conclude that the C-2 group is (stochastically) higher in compliance than the C-1 group.

Pretraining Variations

The conformity behavior of the groups in the first of the pretraining variations on each of the forms of the measure is summarized in Table 1. The *t* for the difference between correlated means on these data was not significant ($p > .05$).

TABLE 1
MEANS AND STANDARD DEVIATIONS OF YIELDING ERRORS ON FORMS A AND B OF THE MEASURE*

Form	*N*	Mean	*SD*
A	80	4.5	3.7
B	80	4.9	3.9

* $t_{corr.} = 1.2$ ($p > .05$).

The reliability of the measure was determined by correlating Form A with Form B. The obtained correlation coefficient, corrected by the Spearman-Brown formula, was .80. These data provide evidence that, for all practical purposes, the two forms are equivalent.

The summary of data providing evidence that the influence effect decreases over a temporal period is shown in Table 2. Since the difference between the forms was not significant, the groups responding first to Form A were combined with those responding first to Form B for this analysis. The scores on the forms of the task administered first were then compared with those administered last. As can be observed in Table 2, the mean score for the test administered second in the sequence is significantly lower ($p < .001$) than that administered first.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF YIELDING ERRORS ON TASKS ADMINISTERED IN THE FIRST AND SECOND HALVES OF THE ITEM SERIES*

Order	*N*	Mean	*SD*
First half	80	5.3	4.1
Second half	80	4.1	3.4

* $t_{corr.} = 3.9$; $p < .001$.

These results are further substantiated by the data, from a replication with other experi-

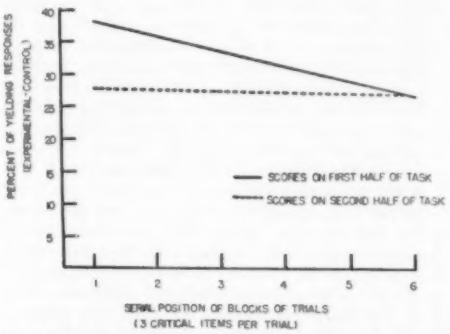

FIG. 1. THE RELATIONSHIP BETWEEN SERIAL POSITION OF ITEMS AND PERCENTAGE OF YIELDING ERRORS

TABLE 3
MEANS AND STANDARD DEVIATIONS OF YIELDING ERRORS OF FIVE GROUPS IN WHICH THE PRETRAINING TASK WAS MANIPULATED

	Experimental Group*	N	Yielding Errors on the Criterion Task	
			M	SD
A	0 critical trials	29	7.56	4.58
B	6 critical trials	27	6.42	4.48
C	12 critical trials	27	5.81	2.83
D	18 critical trials	28	4.82	3.16
E	24 critical trials	28	4.79	3.61

* Groups are described in terms of the number of critical trials on the pretraining task.

mental groups, presented graphically in the smoothed curves of Fig. 1. The percentages of errors on the first half of the test are consistently higher than those on the second half, except in the sixth block of critical trials, supporting the results obtained in the former analysis.

On the basis of these observations it was hypothesized that confidence in the group was related to the frequency of errors made by the majority. This hypothesis was tested by the experimental variation in which the number of critical items on the first half of the measure was varied, while the second half of the measure was retained as the index of conformity.

The results of this experiment are summarized in Table 3 where it is seen that as the number of critical items on the pretraining task is increased, the number of yielding errors decrease. *Ss* responded more accurately (independently) on the second half of the test when the pretraining trials included a greater number of critical trials. Conversely, they conformed to a greater extent when the pretraining task included fewer critical items.

TABLE 4
ANALYSIS OF VARIANCE OF YIELDING ERRORS FOLLOWING PRETRAINING VARIATIONS

Source	df	MS	F
Between groups	4	38.57	2.56*
Within groups	134	15.08	
Total	138		

* $p < .05$.

TABLE 5
MEANS AND STANDARD DEVIATIONS OF YIELDING ERRORS OF MOTIVATED AND CONTROL SUBJECTS*

Condition	N	Conformity Score	
		M	SD
Motivated	30	9.97	6.68
Control (C-3)	25	16.20	7.66

* Variance ratio = 1.31; $p > .10$; $t = 3.22$; $p < .01$.

An analysis of variance of these data is presented in Table 4. In the Bartlett test for homogeneity of variance of these data, the obtained χ^2 of 1.27 was not significant ($p > .05$). The difference between the means of the groups tested, using the simple randomized analysis of variance design, was significant ($p < .05$) permitting rejection of the null hypothesis. The conclusion is that the difference in conformity scores is determined by the number of critical trials on the pretraining task preceding the measurement of conformity on the criterion task.

The Personal Involvement and Evaluation Variations

The data on control *Ss* (C-3), who were instructed that the signals represented the responses of the other members, were compared with those from the experimental groups who were given the additional instruction that the scores on the conformity measure were a reflection of intelligence. These comparisons are summarized in Table 5 where it may be seen that the mean score for the motivated group is significantly lower ($p < .01$) than that of the control (C-3) group. The variance

TABLE 6
MEANS AND STANDARD DEVIATIONS OF YIELDING ERRORS OF *Ss* RECEIVING HIGH AND LOW EVALUATIONS OF PERFORMANCE

Group	N	Errors Before Evaluation		Errors After Evaluation	
		M	SD	M	SD
Recipients of:					
"High" evaluations	25	5.68	3.62	4.88	3.69
"Low" evaluations	25	5.96	3.63	6.76	4.58

TABLE 7
SUMMARY OF ANALYSIS OF COVARIANCE OF YIELDING ERRORS OF EXPERIMENTAL GROUPS RECEIVING HIGH AND LOW EVALUATIONS OF PERFORMANCE

Source	*df*	*MS*	*F*
Adjusted *Ss*	1	32.89	5.30*
Within groups	48	6.20	
Total	49		

* $p < .05$.

ratio for these data is 1.31 and is not significant ($p > .10$).

In the evaluation variation of this experiment, the first half of the series of tasks was used in the analysis as a measure of *Ss*' general dispositional tendency to comply with the judgments of others. The second half of the series of tasks was used to measure the effects of the two different evaluations. This procedure permitted the use of each *S* as his own control in comparing the responses of *Ss* who were given "high" evaluations with those who were given "low" evaluations.

The results of this experiment are summarized in Table 6 in which are presented the means and standard deviations of scores on the pre-evaluation and postevaluation trials. The scores on the pre-evaluation trials (Form A only) are comparable to those of the experimental group used in the personal involvement variation (Table 5). In addition, the scores on the pre-evaluation trials for the "high" evaluation groups are higher than the scores for the postevaluation trials, consistent with the findings reported earlier in this paper. The variances between groups were homogeneous for each series of trials. The t between the two groups on the first series of trials was less than 1.0. Accordingly, the groups were assumed to be drawn from a common population.

The analysis of covariance was employed in the analysis of the data for the postevaluation trials. This analysis is summarized in Table 7. The obtained F value of 5.30 is significant ($p < .05$). This difference cannot be accounted for by a difference in the mean level of initial compliance, since the means of the groups on the experimental trials have been adjusted in the analysis to a common mean initial level of performance on the preliminary trials. The conclusion is that the "high" evaluation of *Ss*' performance resulted in a decrease in the number of yielding errors, while the "low" evaluation resulted in an increase in the number of yielding errors on the postevaluation series.

DISCUSSION

The concern in the present study was with informational social influence, that is, the tendency to uniformity, and the modification of this tendency in the absence of formal social structures. The degree of conformity under these circumstances was hypothesized to be a function of (*a*) the relative confidence the individual has in his own judgment as compared with the judgments of others; and (*b*) his motivational orientation toward the task with which he is confronted.

Confidence in the judgments of others was found to be related to the frequency of errors made by the others. This conclusion is supported by the findings that a general decrease in yielding errors accompanies an increase in the number of trials. Similarly, an increase in the number of "errors" supposedly made by the others on the pretraining task results in a general decrease in the number of yielding errors.

Confidence in the majority was also found to be related to the certainty of the individual that he was receiving information from the others. When the individual responds under "private" conditions there is zero certainty, with no more than chance errors; when he responds under conditions of input without social or communicative value, as in the C-2 control group, the subjective probability is low; and in the experimental groups the subjective probability that information was being received from the others was presumably high. It should be noted that in the second control group (C-2) the schedule of lights was the same as that shown to the experimental groups and thus was not completely arbitrary, i.e., over 50% of the light sequences were correct. Thus, it would appear reasonable to assume that some of the *Ss* were influenced by the "nonarbitrary" sequence of the lights.

Confidence in the individuals' own judgments was found to be related to the evaluations they received of their performance. Thus, the evaluation that *S* was "highest" or "among the highest" of the group, resulted in a decrease

in yielding errors. Conversely the evaluation that *S* was "lowest" or "among the lowest" of the group resulted in an increase in yielding errors. These results supplement the finding that members of a group who change toward the group opinion are more certain of their judgments (Hochbaum, 1954) and indicate that "certainty of judgment" may be subjectively achieved by "copying." The data also support earlier findings that certainty of judgment can be manipulated by success or failure and that negative reinforcement increases convergence toward the group opinion, while positive reinforcement tends to maintain the original level of yielding or independence (Crutchfield, 1955; Kelman, 1950).

The instruction that the individual's performance on the tasks was a reflection of his intelligence resulted in an influence to decrease the number of yielding errors. It would appear that this instruction enhanced the necessity to perform well on the tasks. That is, to perform poorly would be a threat to the individual's self esteem. Knowing nothing about the other persons in the group, the individual tends toward a degree of independent judgment that is maintained if he receives a favorable evaluation of his performance by an "authority," or that is modified toward yielding behavior if he receives an unfavorable evaluation. An alternative explanation is that the statement that "performance on the tasks is related to intelligence" contains the implied instruction that the individual should give an independent judgment rather than rely on the reactions of the others.

SUMMARY

The general hypothesis of this study was that the individual's acceptance of the judgments of others varies as function of: (*a*) the relative confidence he has in his own judgment; and (*b*) his motivational orientation toward the task with which he was confronted. It was assumed that these variables are particularly relevant in the modification of informational social influence.

The degree of conformity was found to: (*a*) increase as the subjective probability that *S* was receiving information from the others increased; and (*b*) decrease as the frequency of errors made by the "majority" increased. Motivational effects induced by the instruction that the task reflected the individual's intelligence resulted in a decrease in yielding. Favorable evaluation of *S*s' performance under this informational influence resulted in a decrease of the level of independence, while unfavorable evaluation resulted in an increase in yielding tendencies.

REFERENCES

CRUTCHFIELD, R. S. Conformity and character. *Amer. Psychologist*, 1955, **10**, 191–198.

DEUTSCH, M., & GERARD, H. A study of normative and informational social influences upon individual judgment. *J. abnorm. soc. Psychol.*, 1955, **51**, 629–636.

DI VESTA, F. J. *Susceptibility to pressures toward uniformity of behavior in social situations: A study of task, motivational and personal factors in conformity behavior.* (ASTIA No. 158–291; AFOSR TR No. 58–70.) Washington, D. C.: Air Force Off. Scientific Res., 1958.

HOCHBAUM, G. M. The relation between group members' self-confidence and their reactions to group pressures to uniformity. *Amer. sociol. Rev.*, 1954, **19**, 678–687.

KELMAN, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. abnorm. soc. Psychol.*, 1950, **45**, 267–285.

NAKAMURA, C. Y. Conformity and problem-solving. *J. abnorm. soc. Psychol.*, 1958, **56**, 315–320.

THIBAUT, J. W., & STRICKLAND, L. Psychological set and social conformity. *J. Pers.*, 1956, **25**, 115–129.

Received June 30, 1958.

LEARNING, RETENTION, AND CONCEPTUAL USAGE IN ELDERLY PATIENTS WITH MEMORY DISORDER[1]

JAMES INGLIS[2]

Institute of Psychiatry, Maudsley Hospital, London

THE investigation reported in this paper was conducted with a threefold purpose. First, it was desired to confirm the principal findings of a previous investigation (Inglis, 1957). These findings were (*a*) that elderly psychiatric patients clinically assessed as suffering from some disorder of memory had great difficulty in the acquisition stage of learning, and (*b*) that such patients also tended to produce a larger mean discrepancy between the Verbal and Performance scales of the Wechsler-Bellevue intelligence test (Wechsler, 1944). In the experimental group, the mean Verbal score was much higher than the mean Performance score than was the case in a group of control patients matched for sex, age, and verbal ability.

Second, it was desired to extend these findings so as to discover if (*a*) the acquisition defect could also be elicited on a learning task mainly manipulative in nature as well as on the mainly verbal, paired-associate learning tasks used in the first investigation, and (*b*) if an impairment of retention could also be demonstrated over and above the acquisition defect.

Third, it was desired to explore certain relationships which appear to exist between learning functions and some other aspects of cognitive ability in elderly psychiatric patients and to discover if such relationships could be brought together within some systematic psychological framework.

These latter tendencies may be traced in at least three studies. Thus, Cameron (1939) initially showed that of a group of six senile patients who gave clinical evidence of serious deterioration, none was able to perform adequately on the Hanfmann-Kasanin (1936) sorting test of conceptual ability. A connection between ability on such conceptual tasks and performance on the Wechsler scale was later demonstrated by Cleveland and Dysinger (1944). They tried to relate certain qualitative aspects of senile deterioration as estimated by performance on the Goldstein (1941) Object Sorting Test to Wechsler results. They showed, in fact, that a low or zero score on the Performance scale of the Wechsler, even when this went together with a much higher score on the Verbal scale, was often accompanied by an inability to handle the type of abstractions necessary for success on the sorting test. Finally, Inglis (1957) demonstrated a relationship between such Wechsler discrepancies and impaired ability in the acquisition phase of paired-associate learning in elderly psychiatric patients.

A psychological system which may be most capable of bringing these findings together is the neuropsychological model propounded by Hebb (1949). Briefly stated, his view has it that adult human learning is mainly a matter of the reorganization of associational processes (i.e., "cell-assemblies" and "phase-sequences") already existing rather than the simple formation of stimulus-response bonds; that is, it is essentially conceptual in character.

One further assumption generates certain expectations which may be tested experimentally. This assumption is that the basic disorder of elderly patients clinically assessed as suffering from memory disorder is an abnormal inability to effect interfacilitations between assemblies and sequences and that this difficulty is inversely related to the number of previous interfacilitations effected and to their recency.

It has been elsewhere suggested (Inglis, 1958) that Hebb's formulations may be able to take into account other observations of cognitive deficit in elderly people, but it would seem sufficient to suggest here two general expectations which follow from the present adaptation of his theory. These are (*a*) that individuals showing impairment on learning tasks should also show a deficit in situations

[1] The author wishes to thank M. B. Shapiro for his helpful comments on this study. He is also indebted to Felix Post for his help and for the facilities he provided.

[2] Now at Queen's University, Kingston, Ontario.

requiring the operation of conceptual processes, such as are commonly supposed to enter into performance on sorting tests, and (*b*) that even the most memory-disordered elderly patients would perform very little differently from control subjects as long as the existence of conceptual linkages could be ensured by maximizing the likely frequency and recency of such links.

METHOD

Subjects

The experimental group comprised 10 elderly patients clinically estimated to be suffering from undoubted memory disorder. This estimation was made by the consultant psychiatrist in charge on the basis of a full psychiatric investigation of each patient. The main considerations entering into his judgment were (*a*) that evidence emerged during the examination that the patient's memory was impaired and (*b*) that the patient's relatives complained about his memory.

A control group of another 10 patients without memory disorder was also tested. Each of these was matched as closely as possible with a member of the experimental group for age and for Wechsler Verbal scale intelligence. The age range of the control group was 62 to 79; that of the experimental group was 65 to 82. The Verbal scale IQ range in the control group was 85 to 103; in the experimental group it was 81 to 106. There were four men and six women in each group, all psychiatric patients of the Geriatric Unit of the Bethlem Royal Hospital. They were predominantly working-class; in each group, four had been in unskilled employment; three had been semiskilled, and three had had skilled jobs. All the patients (both experimental and control) were cooperative during testing, and none of them suffered from any sensory defect so severe as to interfere with test performance. None of them had acted as subjects in any previously reported investigation. The experimental group was mainly diagnosed as cases of senile dementia, usually having been admitted to hospital for difficulties associated with failing memory. The control group was principally diagnosed as affective disorders with such symptoms as depression, agitation, and tension.

Tests

The following tests were administered in an attempt to answer the various problems of confirmation, extension, and exploration outlined above.

Wechsler Adult Intelligence Scale. The WAIS (Wechsler, 1955) was chosen as a measure of intelligence, rather than the Wechsler-Bellevue Form I used in the previous study, for two reasons. First, it is better standardized over the age range of the subjects concerned. Second, a recent paper by Barnes, Busse, & Friedman (1956) showed some differences in the results of elderly patients on the two tests. These authors were uncertain whether such differences were a function of the tests themselves or of the subjects tested. There is some evidence that the patients admitted to the Bethlem Royal Hospital's Geriatric Unit show relatively stable intelligence test scores so that any difference in results might reflect on the generality of the functions tested by the two scales.

Ammons Full Range Picture Vocabulary Test. The Ammons test (Ammons, Larson, & Shearn, 1950) was administered to obtain a measure of verbal ability as far as possible through modalities other than those involved in oral question-and-answer. In this test, the patient simply selects from a multiple-choice set a picture in which a representation of the required word appears. The use of this test also provided an additional check on the question of whether or not the results of the disordered group on certain tests could be referred to perceptual rather than, say, associative or other difficulties.

Paired-Associate Learning Tests. As learning tasks, three sets of three pairs of associates were each given twice as follows: (*a*) In the "new-pair" task the patient had to learn to associate three pairs of words which may be assumed to have in general a low association value (flower–spark; table–river; bottle–comb). The examiner first read the list of words two at a time, at the same time asking the patient to remember which words went together. After this presentation, the examiner read the stimulus words to the patient one by one in random order. The patient then had to recall which response word went with each stimulus word. If the patient could not give the response or gave the wrong one, the correct word was read to him by the examiner. If the patient gave the right answer, it was acknowledged and the next stimulus was given. This learning subtest, like the others, was given twice to secure scores for both acquisition and retention. The acquisition score was the number of trials required by the subject to learn the associates to a criterion of three consecutive correct repetitions of each pair, or twenty repetitions of each pair, whichever came first. The retention scores consisted in a relearning score (or the number of re-presentations of each pair necessary, after a filled interval of about half an hour, to relearn to the criterion of three consecutive correct repetitions of each pair) and a retained members score (or the number of items correctly recalled on the first presentation of the stimuli on the first relearning trial). (*b*) In the "mediate-pair" task, the patient had to learn to associate three pairs of words which it was assumed would have more, and more frequent and recent, conceptual links than the "new-pairs" (cup–plate; cat–milk; gold–lead) but still not maximally frequent and recent associations with each other. The method of administration and the criteria of scoring were exactly as in the previous subtest. (*c*) In the "old-pair" task, the patient had to associate three pairs of words which appeared to have as much likelihood of frequent and recent direct linkage as possible (knife–fork; east–west; hand–foot). The administration and scoring were again the same.

It can be seen that the results of these tests could be expected to bear directly on the second expectation drawn from the extension of Hebb's notions proposed above.

Rey-Davis Test. The Rey-Davis test (Rey, 1934) was

included as a learning task mainly manipulative in character. Its clinical use, mainly with organic patients, has previously been described by Zangwill (1943, 1946) It consists of four 6-in. square boards, each having three rows of three small pegs. Only one peg on each board is fixed, and it is in a different position on each board. The boards are presented in a regular series. In this experiment the subject was told the position of the fixed peg on the first board and had to discover the fixed peg on the other boards by trial and error. Administration was continued either until the patient selected the right peg on each board on two successive series of presentations or until each board had been presented 20 times, whichever came first. This number of presentations was used as a trials score. The successes score consisted of the number of pegs correctly selected on at least two successive occasions.

Shaw Test. The Shaw Test (Bromley, 1955) is a sorting test of conceptual usage. It consists of four wooden blocks which can be arranged in a number of different series (e.g., in terms of height, color, weight, etc.). Bromley (1956) administered the test to 256 normal subjects, including 64 between the ages of 66 and 82. He objectively defined 15 normal abstract sequences which he called A responses, and, in addition, defined statistically 41 less frequent B responses, 16 C responses, and 5 D responses. This index of conceptual usage was selected because, unlike most other sorting tests, it has been standardized on an elderly population, and it has been shown that at least normal elderly people show individual differences on it.

Test Sequence

The tests were usually given over four testing sessions of between about an hour and two hours each. The order of administration was as follows:

Session 1. WAIS Verbal scale
Ammons Vocabulary test
Session 2. "New-pairs" association test—acquisition trials
WAIS Performance scale
"New-pairs"—relearning trials
Session 3. "Mediate-pairs" association test—acquisition trials
Rey-Davis test
"Mediate-pairs"—relearning trials
Session 4. "Old-pairs" association test—acquisition trials
Shaw Test
"Old-pairs"—relearning trials.

The interval between the acquisition and relearning trials in the second session was about half an hour; in the third session, 20 min., and 15 min. in the fourth session. One flaw in the present investigation was that no exact measure was taken of the duration of these intervals to enable strict comparisons to be made. It is unlikely, however, that this was of major importance.

RESULTS AND DISCUSSION

The results obtained may be conveniently considered in terms of the three purposes of the study enunciated above.

Confirmation of previous findings. Comparison of the principal significant variables from the previous study with the equivalent measures used in the present experiment shows a remarkable similarity in the distributions of scores for the two experimental and two control groups (see Table 1). It may be concluded, therefore, that these elderly patients show a striking consistency of mental functioning and that their general cognitive abilities and impairments may be ascertained equally well by

TABLE 1

COMPARISON OF TWO EXPERIMENTAL, AND TWO CONTROL GROUPS IN TERMS OF AGE, WECHSLER WEIGHTED SCORES, AND PAIRED-ASSOCIATE LEARNING TEST SCORES ASSESSED BY MEANS OF STUDENT'S *t*

Variable	Experimental Groups						Control Groups					
	1 (N = 8)		2 (N = 10)		t	p	1 (N = 8)		2 (N = 10)		t	p
	m	σ	m	σ			m	σ	m	σ		
Age	65.75	8.19	70.80	5.67	.49	NS	67.25	7.17	71.50	5.36	.46	NS
Wechsler Weighted Scores												
Verbal scale[a]	33.88	8.90	33.20	9.07	.05	NS	30.50	8.42	34.70	5.62	.40	NS
Performance scale	13.50	13.77	14.10	10.58	.03	NS	26.50	9.75	28.10	6.77	.13	NS
Verbal minus performance	20.38	13.53	19.10	6.28	.08	NS	4.00	9.97	6.60	5.08	.23	NS
Full scale	47.38	18.84	47.30	18.67	.003	NS	57.00	15.25	62.80	11.36	.30	NS
Paired-associates (acquisition score)	59.25	23.60	54.80	28.07	.11	NS	14.13	4.55	10.70	5.16	.47	NS

[a] In each case the Verbal scale weighted score is based on five subtests; i.e., exclusive of the Information subtest.

either the Wechsler-Bellevue or the Wechsler Adult Intelligence scale despite their slight differences in content and scoring.

The means and standard deviations and the significance of the difference between the means of the experimental and control groups in the present study are shown in Table 2. The significance of the differences was estimated by Sandler's (1955) A test. As the method of matched pairs was used, differences in variance between the groups were disregarded in this analysis (Dixon & Massey, 1951).

Previous findings are confirmed by the significance of the Verbal-Performance difference and by a significant difference on the "new-pair" association test acquisition score.

One change in the findings on this occasion was that the criterion groups were significantly different in mean Performance scale score on the WAIS, probably because of a reduction of variance on this score in the two groups. Another was the absence of a significant correlation between the Verbal-Performance discrepancy and the acquisition score on the paired-associate test within the experimental group, such as was shown on the previous occasion. The rank order correlation was then .83 (significant at the .01 level); in this study it dropped to a nonsignificant −.17. The correlation between these two variables within the control group was not significant on either occasion.

Extension of previous findings. The second purpose of this study, it will be recalled, was to extend the previous findings. This was done in a number of ways.

For example, results on the Ammons Vocabulary test show that, as on the WAIS Verbal scale, the criterion groups were not significantly different in terms of mean scores. This finding provides additional evidence to support the view that the difficulties otherwise shown by the memory-disordered group (e.g., on the WAIS Performance scale) cannot be referred to perceptual difficulties alone. The correlation between the Ammons test score and the WAIS Verbal scale IQ within the experimental group proved to be significant only at about the .10 level (rho = .47), but the same correlation within the control group was significant at beyond the .05 level (rho = .70).

TABLE 2
MEANS, STANDARD DEVIATIONS, AND THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN THE MEANS OF THE CRITERION GROUPS ON THE VARIABLES EXAMINED

Variable	Experimental Group		Control Group		A	p
	m	σ	m	σ		
Age	70.80	5.67	71.50	5.36	2.837	NS
WAIS IQs						
Verbal Scale	89.80	8.79	92.50	6.38	1.099	NS
Performance Scale	78.00	12.40	97.00	8.72	0.136	<.001
Verbal minus Performance	11.80	8.15	−4.50	6.47	0.128	<.001
Full Scale	83.80	10.41	94.30	7.17	0.184	.01
Ammons Test	56.40	10.47	58.10	8.54	6.280	NS
"New-pairs"						
Acquisition	54.80	28.07	10.70	5.16	0.135	<.001
Relearning	56.80	34.93	3.90	3.48	0.136	<.001
Retained members	0.20	0.63	1.30	1.16	0.174	.01-.001
"Mediate-pairs"						
Acquisition	26.50	21.81	4.80	1.93	0.189	.02-.01
Relearning	8.70	7.63	1.00	1.25	0.180	.01-.001
Retained members	1.50	1.18	2.30	0.82	0.344	NS
"Old-pairs"						
Acquisition	4.90	1.91	4.00	1.49	0.901	NS
Relearning	1.70	3.95	0.10	0.32	0.664	NS
Retained members	2.60	0.84	2.90	0.32	1.000	NS
Rey-Davis Test						
Trials	20.00	0.00	7.10	5.40	0.116	<.001
Successes	0.90	0.74	3.80	0.63	0.106	<.001
Shaw Test						
A scores	0.30	0.48	2.20	1.03	0.136	<.001
B scores	0.00	0.00	0.60	0.84	0.278	NS
C scores	1.00	2.11	0.60	0.70	2.375	NS
D scores	3.70	1.89	1.10	1.29	0.163	<.001

Results on the Rey-Davis test showed patients in the experimental group to be significantly inferior both in terms of their trials score and their successes score. These results indicate that the kind of acquisition defect elicited in the mainly verbal paired-associate learning type tasks could also be demonstrated on a more purely manipulative test of learning ability.

A further extension of the results of the previous investigation was produced by measures of relearning. Considering for the moment only the "new-pair" variable, it can be seen

from Table 2 that the experimental group was significantly inferior to the control group on both the relearning and the retained members scores. This implies that patients categorized clinically as suffering from memory disorder appear to suffer not only from a defect in the acquisition stage of the learning process but, also from a defect of retention once the material to be learned has been acquired. Since only 3 out of 10 of the experimental group failed completely to reach the criterion of learning in the acquisition trials, this retention difficulty cannot be referred simply to a complete inability to learn at all.

Exploration of an hypothesis. The first expectation to be tested was that individuals showing impairment on learning tasks also show a deficit in situations requiring the operation of such conceptual processes as are entailed in sorting tasks.

The performance of the experimental group on the Shaw test of conceptual ability was inferior to that of the control group in that they obtained significantly fewer A scores (i.e., sortings classifiable as "clear, logical sequences") and significantly more D scores (i.e., sortings classifiable as "illogical and incoherent") according to the criteria set up by Bromley (1956). This finding goes some way towards confirming the hypothesis concerning the relations between conceptual functioning and learning behavior in human adults. It also suggests that the experimental observations made at various times by Cameron (1939), Cleveland and Dysinger (1944), and Inglis (1957) can be drawn together into a kind of psychological syndrome of dysfunctions in which defects of conceptual behavior in the solution of new problems, and in the acquisition and retention of new learning are associated. It appears that the notions proposed by Hebb concerning the relations between cognition and learning may, to some extent at least, provide a suitable framework within which to describe these relationships.

The second hypothesis was that if the factors of frequency and recency could be taken into consideration in such a way as to ensure that conceptual linkages were maximized in the repertoire of even the most disordered, then such patients would perform little differently from their controls.

The tasks used to evaluate this expectation were the three kinds of paired-associate learning test described as "new," "mediate," and "old," respectively. In each case, a measure of acquisition and two measures of retention (i.e., relearning and retained members score) were obtained in the manner described. The results obtained, as can be seen from Table 2, confirmed the expectation.

The experimental group performed very poorly on the "new-pairs," between which it may be assumed that frequent and recent convergences are minimal. The difference between experimental and control subjects diminishes on the "mediate-pairs" and disappears on the "old-pairs," where it may be assumed that such convergences and, hence, meaningfulness, are almost at a maximum.

This assumption is supported by the fact that, within the groups, the new-mediate, and mediate-old differences of the means for both the acquisition and relearning scores are nearly all significant at or beyond the .02 level. The only exception is the mediate-old relearning difference within the control group, which is significant at beyond the .05 level.

SUMMARY

This study was designed to confirm that elderly patients clinically judged to be suffering from memory disorder show difficulty in learning paired-associates and the kind of tasks found in the Wechsler Performance scale. It was also desired to show that the learning difficulty extends to other kinds of learning problems and that retention processes are also affected. The study was also intended to examine the relevance of some of Hebb's notions concerning the likely relations between learning and cognitive functioning in human adults.

Two groups of 10 elderly psychiatric patients, one group of patients with memory disorder and one group without, were tested. The results obtained in general confirmed expectations and suggested that the sort of neuropsychological model elaborated by Hebb may be a useful source of working hypotheses in this area.

REFERENCES

AMMONS, R. B., LARSON, W. L., & SHEARN, C. R. The Full Range Picture Vocabulary Test. V. Results for an adult population. *J. consult. Psychol.*, 1950, **14**, 150–155.

BARNES, R. H., BUSSE, E. W., & FRIEDMAN, E. L. Psychological functioning of aged individuals with normal and abnormal electroencephalograms. II. A study of hospitalized individuals. *J. nerv. ment. Dis.*, 1956, **124**, 585–593.

BROMLEY, D. B. Notes on the Shaw Test. *Brit. J. Psychol.*, 1955, **46**, 310–311.

BROMLEY, D. B. Some experimental tests of the effect of age on creative intellectual output. *J. Geront.*, 1956, **11**, 74–82.

CAMERON, N. Deterioration and regression in schizophrenic thinking. *J. abnorm. soc. Psychol.*, 1939, **34**, 265–270.

CLEVELAND, S., & DYSINGER, D. Mental deterioration in senile psychosis. *J. abnorm. soc. Psychol.*, 1944, **39**, 368–372.

DIXON, W. J., & MASSEY, E. J. *Introduction to statistical analysis*. New York: McGraw-Hill, 1951.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior; an experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 2 (Whole No. 239).

HANFMANN, EUGENIA, & KASANIN, V. A method for the study of concept formation. *J. Psychol.*, 1936, **3**, 521–540.

HEBB, D. O. *The organization of behavior. A neuropsychological theory*. New York: Wiley, 1949.

INGLIS, J. An experimental study of learning and memory function in elderly psychiatric patients. *J. ment. Sci.*, 1957, **103**, 796–803.

INGLIS, J. Psychological investigations of cognitive deficit in elderly psychiatric patients. *Psychol. Bull.*, 1958, **55**, 197–214.

REY, A. D'un procédé pour évaluer l'éducabilité. *Arch. Psychol. Geneve*, 1934, **24**, 297–337.

SANDLER, J. A test of the significance of the difference between the means of correlated measures based on a simplification of Student's *t*. *Brit. J. Psychol.*, 1955, **46**, 225–226.

WECHSLER, D. *The measurement of adult intelligence*. (3rd ed.) Baltimore: Williams & Wilkins, 1944.

WECHSLER, D. *Wechsler Adult Intelligence Scale Manual*. New York: Psychological Corp., 1955.

ZANGWILL, O. L. Clinical tests of memory impairment. *Proc. Roy. Soc. Med.*, 1943, **36**, 576–580.

ZANGWILL, O. L. Some clinical applications of the Rey-Davis performance test. *J. ment. Sci.*, 1946, **92**, 19–34.

Received July 7, 1958.

A COMPARATIVE STUDY OF THE PERSONAL HISTORIES OF SCHIZOPHRENIC AND NONPSYCHIATRIC PATIENTS[1]

WILLIAM SCHOFIELD AND LUCY BALIAN

University of Minnesota

ONE of the currently popular views of the etiology of severe personality disruption holds that the seeds of mental illness are to be found in the life experiences of the individual. In particular, the childhood is conceived as a critical period, and certain areas, such as parent-child relationships and psychosexual development, are viewed as crucial determiners of or prodromal to later adjustment or psychopathology. Three possible forms of "historical" or biographic etiology may be distinguished: (*a*) the traumatic incident (e.g., witnessing in childhood the "primal scene"), (*b*) sequential traumata (i.e., a concatenation of emotional blows resulting finally in disintegration of the ego), and (*c*) acquired predisposition (i.e., the learning of a pattern of maladaptive response). Writing in 1893, Breuer and Freud (trans. 1957) noted that:

> The causal relation between the determining psychical trauma and the hysterical phenomenon is not of a kind implying that the trauma merely acts like an "agent provocateur" in releasing the symptom, which thereafter leads a separate existence. We must presume rather that the psychical trauma—or more precisely the memory of the trauma—acts like a foreign body which long after its entry must continue to be regarded as an agent that is still at work (p. 6).

Freud (Breuer & Freud, trans. 1957), in describing the case of Elizabeth von R., observes:

> Almost invariably when I have observed the determinants of such conditions what I have come upon has not been a *single* traumatic cause but a group of similar ones. . . . In some of these instances it could be established that the symptom in question had already appeared for a short time after the first trauma and had then passed off, till it was brought on again and stabilized by a succeeding trauma (p. 173).

Finally, we have a succinct expression of the learning hypothesis in Shaffer and Shoben (1956):

> An evident conclusion is that our distinctive attributes, even our most fundamentally human qualities, are products of our experience with other human beings. *Personality* (underlining ours) is learned as a result of the events in one's history (p. 402).

The significance of the once-occurring trauma in production of later symptomatology is given less prominence now than it held during the heyday of "la grande hysterie" and when public-compulsive symptoms were viewed as more circumscribed phenomena. Likewise, a simple additive notion about accumulations of emotional "shocks" is retained today chiefly by the laity, who express this notion in the "straw and camel's back" allegory. It is currently more common to view the patient's history as either having deprived him of an opportunity to learn normal patterns of socialization or as "overtraining" him in some repertoire of abnormal responses. If varying patterns of psychopathology are differentiated, and if the biography, or certain ages or elements in the biography, are believed to have pathogenic potential, it is reasonable to assume that differential patterns of personal history might be found which are demonstrably associated with various behavioral syndromes. Research in this realm immediately involves the investigator in difficult methodological problems with respect to definition and recording of the life history.

What is a life history? As the concept is widely and loosely used in much psychological writing, it is implicitly a theoretical abstraction. It may be conceived as inclusive of all *events* in the total sequence of time-space displacements of the individual from the moment of his birth up to some time at which a summary is prepared. It may be conceived as the complete sequence of *experiences* had by the individual over such a period. It may be thought of as a "complete" collection of the events *and* experiences of the individual. Unless one wishes to assume perfect recording (and retention?) characteristics for the nervous apparatus, it is unlikely that the individual's

[1] This study was made possible by a grant from the Frederick L. Wells Fund.

population of events, as observable time-space dispositions, has isomorphic representation in his population of experiences or subjective events.

Without much formal attention to the theory of the life history, most of the research into biographical factors in mental illness has been content to use the framework or part of the structure of the so-called "psychiatric history." This is essentially a relatively uniform selection of significant events and experiences in the histories of research subjects. What determines the selection and what defines significance? Selection is determined in part by what data are obtainable as a function of accuracy of records and integrity of memory. Significance is expressed through clinical judgment, largely as a result of observations of apparent association between certain factors and specific outcomes (e.g., broken homes and delinquency). Significance may be also determined on the grounds of currently (and locally) popular theory. Recognizing these constraints upon the adequacy of the typical psychiatric history as a representative sampling from the theoretical population of events and experiences constituting the true life history, it is incumbent upon the investigator into the contribution of life history factors to psychogenesis of emotional disturbance to seek at least minimal insurance against these sources of error. Such safeguards are afforded in the utilization of history outlines and recording forms and in the collection of data for control samples. The development and application of detailed history schedules helps to assure uniformity and thoroughness of coverage. The study of control groups serves as a check on the associations and hypotheses of etiological factors suggested by the study of exclusively pathological samples.

The obviousness of these experimental caveats encourages an expectation that they have been well respected in researches into the life history as a source of psychopathology. A search of the literature reveals over 300 studies of life histories of psychiatric patients. Fewer than ten of these included data on a reasonably comparable control group (Aldrich & Coffin, 1948; Lane, 1955; Nielsen, 1954; Steinberg & Wittman, 1943). The bulk of the studies report simply the frequency of such items as patient's age at death of a parent as recorded in mental hospital records. Frequently, only selected items are tabulated without any attempt at a comprehensive history. When more extensive coverage of history data has been undertaken, a variety of procedures for collection has been used and efforts to study a control sample have been rare. When studied, control groups have seemed mostly to be determined by accessibility rather than appropriateness (Brockway, Gleser, Winakur, & Ulett, 1954).

PURPOSE AND PREMISES

The purpose of this study was (*a*) to determine what the life histories of "normal" persons look like if examined through the spectroscope of a comprehensive psychiatric history interview as conducted and recorded by a skilled clinician and (*b*) to determine in what ways such histories may be distinguishable from those of such psychiatric patients as schizophrenics.

The basic hypothesis of this research simply stated was: the life histories of "normal," i.e., nonpsychiatric patients, will be readily and clearly differentiated from those of psychiatric patients if equally complete and carefully collected history data are available for both. More specifically, the hypothesis is that the histories of normals will reveal markedly less occurrence of those events or experiences (trauma, deprivations, frustrations, conflicts, etc.) which have been commonly considered to have a psychogenic or prodromal role in schizophrenia. This is a formal statement of the assumption which is made, usually implicitly, in those clinical and uncontrolled studies which derive from the absolute frequency of given events in the histories of certain patients the conclusion that the event has causal import. In the absence of comparative data on the frequency of these same events in appropriate control subjects, there is not only room but need to be doubtful of such conclusions.

SAMPLE AND PROCEDURE

The selection of the "normal" subjects (*S*s) for this study was determined primarily by the nature of the 178 schizophrenics with whom

comparison was to be made. These patients were hospitalized at the University of Minnesota Hospitals, and comprehensive personal history statistics have been previously reported (Schofield, Hathaway, Hastings, & Bell, 1954) for them.[2] To assure comparability of the normal and psychiatric samples, the former were drawn primarily from the same population which yielded the latter, i.e., the general group of persons referred to the University hospitals for diagnosis and treatment. Of the total, 105, or 70% were University patients. An additional 14 cases were obtained from Minneapolis General Hospital. Finally, 31 physically and psychiatrically negative cases were obtained from sources including hospital employees, students, employees of a large industrial firm, and office workers. The term "normal" as applied to this sample specifies the absence of psychiatric disturbances under treatment at the time the histories were recorded *and* the absence of any previous mental disorder. The nonpsychiatric patients were drawn from all but the psychiatric wards and psychiatric clinics of both the University hospitals and Minneapolis General Hospital. The medical diagnoses represented covered a wide range, and no type of illness or defect was predominant.

To assure further comparability of the two groups, selection of the normals was made so as to achieve matching with the schizophrenics for age, sex, and marital status. Success of the matching is indicated in Table 1. As a further reflection of the comparability of the two samples, data on education, rural–urban origin, and religious affiliation are recorded in Table 1. The mean number of years of formal education for the normals is reliably higher than that of the schizophrenics; however, the overlap between the two distributions is approximately 50%. Also, the higher average educational level of the normals undoubtedly reflects the general increment in average educational level of the public, since they were hospitalized in 1956–1957 (of school age in 1904–1956), in contrast to the schizophrenics who were hospitalized between 1938–1944 (of school age in 1886–1944). No basic difference is indicated in the intellectual or general socioeconomic character of the two samples. The comparability of the two samples is further supported by the data on number of sibs and number of children in the married *S*s as reported in Table 1.

TABLE 1
GROSS DESCRIPTIVE DATA FOR NORMALS AND SCHIZOPHRENICS

	Normal (N = 150)	Schizophrenic (N = 178)
Age	M = 28.6 yrs. SD = 9.7	M = 28.7 yrs. SD = 9.6
Sex		
Male	44.7%	44.4%
Female	55.3	55.6
Marital Status		
Married	33.3%	26.9%
Single	64.7	69.1
Divorced	1.3	1.7
Widowed	0.6	1.7
Separated	0.0	0.6
Education	M = 12.0 yrs. SD = 3.4	M = 9.9 yrs. SD = 2.5
Home		
Urban	44.7%	50.8%
Rural	55.3	49.2
Religion		
Lutheran	38.0%	35.7%
Other Prot.	31.7	31.0
Catholic	24.0	27.0
Jewish	0.6	6.2
Moslem	0.6	0.0
No. of Siblings	M = 4.1 SD = 2.8	M = 4.2 SD = 3.4
No. of Children	(N = 53) M = 2.4 SD = 2.1	(N = 50) M = 2.8 SD = 2.4

The life histories of the normals were collected through the medium of a comprehensive clinical interview requiring from 45 to 90 minutes.[3] Only two of the patients who were approached refused to cooperate. Satisfactory rapport was achieved in most cases, and the *S*s appeared to give reliable accounts of their backgrounds. Many of them were interested in the nature of the study, and this was briefly discussed with them at the close of the interview. Also, at the end of the interview each *S* was administered an MMPI as a further check on his psychiatric status.

Immediately after each interview, the interviewer transcribed her notes and further observations onto a detailed schedule covering over 100 distinct items pertaining to developmental, personal, social, and medical history.

[2] It is recommended that this report be read for further detail on the history schedule and method of data collection for the schizophrenic sample.

[3] All interviews were conducted by the junior author, who at the time was a senior medical student.

This was the same schedule which had been used in recording the history data for the schizophrenic sample.[4]

RESULTS

The major findings of the study are reported in Tables 2–7, which show the percentage frequency of occurrence in the two samples of various psychological relationships and adjustment variables. These range from the quality of the relationship between the *S*s parents to the degree of manifestation of a life plan and initiative in the pursuit of that plan. The reliability of the differences between the distributions of the two samples are reported in the tables.[5] Generally more notable than the presence or absence of statistical reliability of the differences are the marked *overlaps* in the distributions for the schizophrenic and normal samples.

Table 2 reveals that the relationship between the parents of *both* the normal *S*s and schizophrenic patients was predominantly one of affection. Relationships characterized as ambivalent, indifferent, or hostile were slightly more frequent between the parents of the normals. In their relationships with their fathers, the schizophrenics experienced a reliably higher frequency of unfavorable attitudes. However, two-thirds of these patients apparently received affection from their fathers, and less than one-fourth were the recipients of either rejection or domination. The maternal relationship more clearly differentiates the two samples. Again, however, it is to be noted that nearly two-thirds of the schizophrenics enjoyed the affection of their mothers. Of the various undesirable relationships which were recorded, overprotection and domination were most prominent in the schizophrenics. This is slight support for the

TABLE 2

INTERPERSONAL RELATIONSHIPS, HOME FACTORS, AND SCHOOL ADJUSTMENT IN THE EARLY HISTORIES OF NORMALS AND SCHIZOPHRENICS

	Normal	Schizophrenic	χ^2	*P*
Interparental Relationship	(*N* = 144)	(*N* = 101)		
Affection	75.7%	76.2%	0.14	>0.98
Ambivalence	6.3	5.9		
Indifference	0.7	1.9		
Hostility	17.4	15.8		
Paternal Relationship	(*N* = 143)	(*N* = 127)		
Affection	76.2%	65.3%	13.02	<.05
Ambivalence	9.8	5.5		
Indifference	4.2	3.1		
Rejection	7.7	11.8	3.91	<.05
Overprotection	0.7	1.5		
Domination	1.4	11.8		
Maternal Relationship	(*N* = 150)	(*N* = 128)		
Affection	81.3%	64.8%	32.90	<.01
Ambivalence	8.0	2.3		
Indifference	2.7	1.5		
Rejection	6.0	6.2	9.72	<.01
Overprotection	0.7	13.2		
Domination	0.7	10.9		
Neglect	0.7	0.7		
Sibling Relationship	(*N* = 178)	(*N* = 126)		
None	3.4%	3.9%	4.43	<.50
Affection	70.8	76.1		
Indifference	2.2	4.7		
Rivalry	16.3	13.4		
Domination	2.8	0.7		
Submission	4.5	0.7		
Home Conditions				
Poverty	20.7%	9.0%	8.00	<.01
Alcoholism	6.0	6.8	0.19	<.70
Invalid	12.7	0.6	18.67	<.01
Divorce	6.0	1.6	3.14	<.10
Separation	2.7	1.1	0.38	<.70
Death of parent	14.7	10.7	0.82	<.50
School Acceptance	(*N* = 155)	(*N* = 159)		
Marked hostility	7.1%	3.8%	13.64	<.01
Mild dislike	17.4	11.3		
Indifference	11.0	25.8		
Agreeable	54.8	54.7		
Keen enjoyment	10.0	4.4		
School Achievement	(*N* = 150)	(*N* = 163)		
Repeated failure	7.3%	4.9%	21.49	<.01
Work difficult	6.0	19.6		
Average performance	42.0	50.2		
Easily earned good grades	38.7	19.6		
Accelerated	6.0	4.4		
School Deportment				
Poor record	9.3%	7.9%	22.92	<.01
Usual no. of escapades	36.0	10.6		
Excellent record	54.7	81.6		

Note.—In this and subsequent tables, where two χ^2 values appear for a given factor, one is based on all categories and the second is for comparison of bracketed versus unbracketed categories.

[4] Copies of the 12-page rating schedule for case history analysis and of the 4-page record form used with it have been deposited with the American Documentation Institute. Order Document No. 4209 from ADI Auxiliary Publications Project, Photo-duplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for 6 x 8 inch photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

This schedule was originally developed by Dahlstrom (1949) in connection with an unpublished doctoral dissertation.

[5] The chi square test was used to determine probability that the obtained distributions belong to a common population. Where indicated, Yates' correction was applied.

TABLE 3
OCCUPATIONAL HISTORIES AND RELIGIOUS ORIENTATION OF NORMALS AND SCHIZOPHRENICS

	Normal	Schizophrenic	χ^2	P
Occupation	(N = 150)	(N = 177)		
None	15.3%	21.1%	0.72	<.50
Occupational Success	(N = 127)	(N = 127)		
Poor	0.0%	14.9%	1.60	<.30
Average	38.6	31.5		
Good	61.4	53.5	18.66	<.01
Occupational Satisfaction	(N = 127)	(N = 126)		
Dislike	7.9%	9.5%	21.74	<.01
Indifference	17.3	43.6		
Enjoyment	74.8	46.8		
Church Attendance	(N = 150)	(N = 147)		
Very infrequent	22.0%	38.3%	4.77	<.10
Occasionally	36.7	31.2		
Steady	41.3	35.3		
Religiousness	(N = 142)	(N = 121)		
Intellectualized, ritualistic	18.3%	4.1%	12.31	<.01
Occasional solace	42.3	55.4		
Dominant source of balance	39.4	40.5		

current belief of some experts in the existence of the so-called "schizophrenogenic mother":

> Psychoanalytic students are uniform in the opinion that maternal rejection and domination are regularly found in the histories of those who are found later to have been predisposed to the development of schizophrenia. The mother of the schizophrenic is variously described as cold, dominating, narcissistic, lacking love for the child, having death wishes toward it . . . (Woolley, 1953).

However, domination or overprotection characterized the relationship of *less than one-fourth* of the mothers of the schizophrenics in the present study.

Table 2 also reports the relationships between the two samples and their respective sibs and the frequency of six factors of physical and/or psychological deprivation or trauma. The intersib relationships do not differentiate the two groups. While none of the home conditions noted occurred in more than a fifth of the homes of either sample, two of the factors, namely, poverty and invalidism, did have a reliably different rate in the two groups, and the rate of divorce approached a reliable difference. All three of these more frequently characterized the childhood homes of the normals.

The attitudes, adjustment, and achievement of the two groups in their school experiences are recorded in Table 2. All three factors show reliably different distributions for the two samples. General attitude toward the school situation was less different for the normals and schizophrenics than were achievement and deportment. Three times as many schizophrenics as normals found their school work difficult, and twice as many normals as schizophrenics easily earned good grades. The passivity of the prepsychotic schizophrenic in the school room, which has been commonly observed by clinicians, is suggested in their clear preponderance of excellent deportment records. While nearly one-fourth of the schizophrenics experienced failure or difficulty with their school work, less than one-tenth of them had poor deportment. By contrast, while better than 85% of the normals had satisfactory or superior achievement, only half of them had excellent deportment records.

Occupational success and satisfaction distinguishes the normals from the schizophrenics as revealed in Table 3. The proportion of the two groups without any history of occupation is not reliably different. While poor occupational achievement was recorded for none of the normals with an occupational history, this rating was assigned to one-fifth of the schizophrenics. Overlap is again notable, with 85% of the schizophrenics manifesting average or good occupational success. Satisfaction with their occupations also differentiated the two groups. Over half of the schizophrenics disliked or were indifferent to their work, while three-fourths of the normals apparently enjoyed their occupations.

Table 3 indicates that the two groups were not differentiated with regard to frequency of church attendance, although the role of religion or attitude toward it was different in the two. While there was no difference in the frequency with which religion afforded as dominant source of balance to the lives of the two groups, an intellectualized or ritualistic approach to religion was found four times as frequently among the normals as among the schizophrenics. Table 3 also reports the occur-

rence of delinquency and criminal records; the rates are very small and not different for the two samples.

Table 4 records the dating history and marital adjustment of the two samples. Frequency of dating is not reliably different in the two groups if the "average" and "very popular" categories are combined for contrast with the "none or little" category. Although there were no schizophrenics rated as "very popular" in terms of frequency of dating, less than 5% of the normals fell into this category. Approximately a third of both samples were married (Table 1). The marital adjustment of the two groups was not different, and a third of both groups experienced frustration or conflict in their marriages. Although the normals tended to derive distinct pleasure from their marriages with somewhat higher frequency than the schizophrenics, a third of both groups evinced attitudes of affection toward their spouses.

The quality of sexual adjustment and adequacy of outlet are shown in Table 4. Sexual adjustment is reliably different in the two samples. Surprisingly, the difference appears primarily in the greater frequency of poor sexual adjustment in the normal *S*s. Likewise, the schizophrenics were rated twice as frequently as the normals as enjoying "good sexual oulets."

TABLE 4

DATING HISTORY, MARITAL AND HETEROSEXUAL ADJUSTMENTS OF NORMALS AND SCHIZOPHRENICS

	Normal	Schizophrenic	χ^2	*P*
Dating	(*N* = 150)	(*N* = 145)		
None or little	49.3%	57.2%	5.98	<.10
Average	46.0	42.8	1.83	<.20
Very popular	4.7	0.0		
Marital Adjustment	(*N* = 53)	(*N* = 50)		
Extreme frustration	18.9%	16.0%	6.73	<.20
Continual conflict	15.1	14.0		
Compatibility	17.0	42.0		
Pleasure	41.5	24.0		
Chief pleasure	7.5	4.0		
Affection to Mate	(*N* = 53)	(*N* = 51)		
	62.3%	64.7%		
Heterosexual Adjustment	(*N* = 150)	(*N* = 171)		
Poor	36.7%	22.2%	10.49	<.01
Fair	20.7	33.3		
Good	42.7	44.4		
Adequacy of Outlet	(*N* = 150)	(*N* = 170)		
Poor	38.6%	24.7%	11.18	<.01
Fair	33.3	30.5		
Good	28.0	44.7		

TABLE 5

SOCIAL SKILL AND ADJUSTMENT OF SCHIZOPHRENICS AND NORMALS

	Normal	Schizophrenic	χ^2	*P*
Social Adjustment	(*N* = 150)	(*N* = 85)		
Withdrawal	10.0%	61.2%	78.60	<.01
Ambivalence	7.3	14.1		
Membership	82.7	24.7		
Social Intelligence	(*N* = 150)	(*N* = 167)		
Inept and clumsy	8.7%	8.9%	29.66	<.01
Moderate skill	68.7	88.8		
Adept	22.7	2.3		
Poise	(*N* = 150)	(*N* = 170)		
Retiring and sensitive	22.7%	37.4%	12.99	<.01
Fairly articulate	51.3	49.7		
Confident	26.0	12.8		
Recreation	(*N* = 150)	(*N* = 163)		
Solitary	27.3%	40.5%	6.51	<.05
Mixed	66.0	52.2		
Social	6.7	7.3		

The quality of social adjustment and of factors affecting interpersonal relationships are shown in Table 5. The adequacy of social adjustment is clearly different for the two samples, with the schizophrenics showing a high frequency of withdrawal and a low rate of active group membership. Likewise, both the variables of social intelligence and poise show reliably inferior distributions for the schizophrenics. The overlaps between the two groups are considerable, however. Nearly 90% of the schizophrenics were rated as having at least moderate skill in interpersonal relations, and only a third of them were characterized as retiring and sensitive. Pursuit of solitary recreation was more caharacteristic of the schizophrenics than the normals, but over half of the former had a history of mixed or social recreation.

In Table 6 are reported five areas of early personal attitude and expression which might be broadly classed as manifestations of individual perspective and morale. Four of these variables show reliably different distributions in the two groups. Summarizing for these variables, it may be said that the schizo-

TABLE 6

INTERESTS, ASPIRATIONS, AND INITIATIVE OF NORMALS AND SCHIZOPHRENICS

	Normal	Schizophrenic	χ^2	P
Breadth of Interest	(N = 150)	(N = 173)		
Narrow	20.7%	35.2%	44.30	<.01
Some outside	46.0	60.2		
Broad	33.3	4.6		
Level of Aspiration	(N = 150)	(N = 170)		
Limited	16.0%	34.7%	17.79	<.01
Interested in improving	61.3	55.3		
High	22.7	10.0		
Life Plan	(N = 150)	(N = 164)		
Vague	22.7%	46.4%	56.63	<.01
Confused	9.3	28.0		
Clear	68.0	25.6		
Stability	(N = 150)	(N = 172)		
Constant fluctuation	10.7%	13.9%	47.17	<.01
Moderate variability	78.7	42.7		
Stolid	10.7	43.4		
Initiative	(N = 150)	(N = 169)		
Apathetic	8.0%	16.0%	4.12	<.20
Appropriate	57.3	53.2		
Energetic	34.7	30.8		

phrenics less frequently had broad interests, a high level of aspiration, and a clear life plan; more frequently than the normals they were characterized by a stolid, nonvarying temperament and by absence of initiative.

DISCUSSION

The single most impressive feature of the data presented in Tables 2-6 is the sizable overlap of the normal *S*s and schizophrenic patients in the distributions of the various personal history variables. Of the 35 separate tests which were run, 13 (or 37%) failed to reveal a reliable difference between the two samples. Further, on 5 of the remaining 22 variables, the distributions showed a reliably greater presence in the normals of negative or undesirable conditions. In those instances where the statistical tests did indicate a reliable characterization of the schizophrenics by prevalence of a pathogenic variable, the normals generally also showed a closely approximating degree of the same factor. Before discussing the implications of the findings, it will be well to review facets of the data collection and recording which might have had biasing effects.

The history data for the schizophrenics were abstracted from the material routinely collected and recorded clinically in the hospital charts of these patients as they were admitted, diagnosed, treated, and discharged. A variety of persons, including social workers, junior medical students, psychiatric residents, and staff psychiatrists, contributed to the recording of these data. No single person was charged with the collection of comprehensive history nor was a detailed research schedule applied as a reference in collecting the material. These facts suggest the possible underestimation of the actual frequency of certain variables in the histories of the schizophrenics. However, in abstracting the clinical material and recording it on the research schedule, no attempt was made to force the rating of a variable when clear information was not available; free use was made of an "unknown" category. This tactic has the effect of enhancing the reliability of the frequencies reported at the expense of having a varying size of sample (for example, "paternal relationships" was rated for only 127 of the 178 schizophrenics).

Reliability of the ratings of the various factors is an important consideration. This is especially true with respect to the data on the schizophrenics for whom the original clinical records were not uniform and from which, without any other source of information or contact with the patient, the rater had to abstract the material pertinent to a given variable and then assign it a rating. The definitions of the various scales were made as objective and nonambiguous as possible, and the number of steps to each scale was generally small. As reported previously, independent abstractors-and-raters achieved an 80–95% agreement over a small sample of trial cases (Schofield et al., 1954).

The reliability of the data for the normals was enhanced by providing for an immediate recording and rating of information which had been collected in an extended interview conducted with the research schedule as an implicit guide to insure coverage. Failure to find distinguishing features for the two groups might be the result of a "contamination effect" if the same person or persons had been responsible for the study and rating of both the schizophreics and normals. Actually, two researchers working quite independently and

at different periods of time collected and rated the schizophrenic and normal material respectively.[6] This avoids an artificial overlap in the distributions as a function of a common rater projecting implicit "base rate" standards from one sample to another. There remains the question of interrater reliability and the possibility of stable but different interpretations of the criteria for various ratings leading to Type I errors (Walker & Lev, 1953). This possibility was particularly suggested, for example, by the surprising distributions of the "heterosexual" adjustment variable (see Table 4). As a check on the possibility that relatively different criteria had been applied by the two raters (both single females) in assessing this variable, they were asked to give independent, free accounts of their respective interpretations of the steps on this scale and of the criteria they utilized in assigning the cases. They appeared to be in essential agreement in these respects. In short, within the limits of the inequalities imposed by the differences in the nature of the raw data for the two samples and the methods by which the basic history material was obtained, it would seem that the ratings of the various history factors were reasonably reliable and no obviously biasing and unbalanced factor operated which would serve to either exaggerate or diminish true differences between the two groups.

Consideration must also be given to the possibility that such differences as were obtained between the two samples might be a function of the different periods of time from which they were sampled. As pointed out in discussing the difference in mean educational level of the two groups, the schizophrenic patients were hospitalized between 1938–1944. Approximately 80% of these patients were between the ages of 20–50 when hospitalized. Using these ages as a reference point for the sample and defining childhood and early adolescence as encompassing the first 15 years of life, this period occurred for the bulk of the schizophrenics between 1890–1940. By contrast, the normals (with the same age distributions as the schizophrenics) were evaluated in 1956–1957. For these *S*s, the period of childhood and early adolescence would fall predominantly in the years 1910–1950. It would be difficult to ascertain the degree to which these two periods would be characterized by distinctive patterns of parental attitude, child-rearing practices, major social upheaval, and other potential sources of psychological effects in the early life histories of individuals. Such contrast could be best drawn for the earliest period represented in the schizophrenics (1890–1910) and the most recent period sampled in the normals (1940–1950), but these periods would account for the "formative years" of only a small portion of each of the groups. The overlap in childhood years for the two samples is great, and such significant factors as the impact of Freudian psychology, the first World War, and the Great Depression occurred in the time interval common to both. It seems unlikely that broad differences in the sociopsychological cultures from which the schizophrenics and normals were drawn can be used to account for their respective distributions on the personal history variables analyzed in this study. In any event, if significant social and cultural factors did indeed differentiate the periods 1890–1910 and 1940–1950, and such factors had causal potency for personality development, they should contribute to more and larger differences between the schizophrenics and normals. The restricted number of differences obtained and the impressive amount of overlap between the two groups suggests limited existence and/or potency of differential sociological factors.

Finally, the lack of more extensive and clear-cut differences between the backgrounds of the two samples might be attributed to the fact that they did not actually represent distinct populations of the psychiatrically ill and the psychiatrically negative. No quarrel can be made with the diagnoses of the schizophrenics. They clearly suffered psychotic disturbance of sufficient magnitude to necessitate hospitalization. Furthermore, they received their specific diagnoses at a time when schizophrenia was not being used as a synonym for all psychosis without obvious brain pathology.

Some clinicians undoubtedly would take exception to the "normality" of our control group. It should be iterated that they denied any history of mental illness or psychiatric consultation, and they manifested no evidence of gross emotional disturbance at the time they were interviewed in spite of the fact that they

[6] The schizophrenic records were rated by Miss Bell in 1951–52, and the normal *S*s were recorded by Lucy Balian in 1956–57.

TABLE 7
MMPI SCORES OBTAINED BY NORMAL CONTROLS

Scale	Mean T Scores				t Ratio
	Female (N = 75)	Male (N = 57)	"Traumatized" (N = 37)	"Nontraumatized" (N = 95)	
?	50	50			
L	53	50	4.27 ± 1.98	4.01 ± 2.21	0.6
F	50	53	3.78 ± 2.98	2.88 ± 2.13	1.66
K	59	59	14.43 ± 5.47	17.56 ± 4.00	4.35*
Hs	56	57	57.73 ± 9.41	55.88 ± 10.30	1.0
D	57	58	62.27 ± 11.11	54.38 ± 9.74	3.75*
Hy	59	60	59.95 ± 8.21	59.26 ± 8.43	0.43
Pd	55	60	57.46 ± 11.53	57.69 ± 9.08	0.11
Mf	51	55	51.76 ± 10.21	52.46 ± 8.55	0.37
Pa	56	53	54.43 ± 7.87	53.77 ± 7.93	0.44
Pt	55	56	56.62 ± 8.86	55.48 ± 8.61	0.67
Sc	55	57	56.81 ± 9.64	55.83 ± 6.86	0.56
Ma	53	55	53.68 ± 9.64	53.94 ± 10.41	0.13

* Exceeds the .01 level.

were currently under study or treatment for serious physical illnesses. The MMPIs administered as an objective check on psychiatric status supported the clinical impression of essential normality. As shown in Table 7, the mean scores on this psychiatric screening test were well within normal limits. A further analysis of the quality of the normal group was made by identifying those *S*s whose histories included one or more of the events or experiences which are generally regarded as psychic trauma. The MMPIs of this group of 37 persons (nearly one-fourth of the normal sample) were compared with those of the remaining "nontraumatized" sample. Mean profiles of the two subgroups are reported in Table 7. Only the *K* and *D* scales distinguish these cases. The "traumatized" group had a lower mean *K* score ($t = 4.35$; $P < .01$) and a higher mean *D* score ($t = 3.75$; $P < .01$). These observations suggest a less defensive and somewhat more depressive orientation in the *S*s with the traumatic histories. The lack of more extensive differentiation of these two subgroups throws a further doubt on the hypothesis that early trauma per se are significant predisposing factors in later mental illness.

The data of this study seem to cast serious doubt on the etiological significance of certain early life factors for which such import has been frequently claimed. These factors may in fact play a causal role in the development of personality disturbance, but not as solitary pathogenic elements. It would appear that it is the patterning or chaining of experiences rather than occurrence or absence which must be examined. While the notion of multiple causation is well established, it is more frequently stated in the context of types of etiological agent—the physical and the psychological—rather than in terms of multiplicity within a given area—personal relationships.

The surprising frequency with which certain forms of pathogenic experiences or circumstances were found in the life histories of the normal *S*s suggests the need to think in terms of "suppressor" experiences or control variables in the development of personality. Woolley (1953) found cold, rejecting, or dominating and exploiting mothers to be "regularly" present in 100 cases selected only for the adequacy of their histories, but stipulated:

> These factors constitute the background for the children who escape as well as for their schizophrenic siblings. Moreover, there are families in which no schizophrenic denouements occur. Evidently, there must be factors concerning the degree of rejection, its time of occurrence, its differential distribution among the siblings or the occurrence of reinforcing or ameliorating experiences.

May it not be that the development of serious mental disorder will be less well understood if we concentrate solely on examination of pathological processes and injurious agents, rather than examining for the nature and extent of "immunizing" experiences? It seems necessary that we turn some of our research energies toward a discovery of those circumstances or experiences of life which either contribute directly to mental health and emotional stability or which serve to delimit or erase the effects of pathogenic events. For this purpose, we will need to make extensive psychological study of the biographies of normal persons as well as of patients, with such biographies recorded so that their coverage and uniformity facilitate analysis.

SUMMARY AND CONCLUSIONS

Through the medium of extended clinical interviews the life histories of 150 psychiatrically normal subjects were collected and subse-

quently recorded in detail on a research schedule which had been used previously in a study of the histories of 178 hospitalized schizophrenics. Of the 150 normals, 119 were hospital or clinic patients being studied and treated for a wide range of serious physical illnesses. Selection of the normals was made so that they were drawn from the same general population as the psychiatric patients, and they were matched with the schizophrenics for age, sex, and marital status.

Separate statistical analyses were made of the reliability of the differences between the distributions of the normals and schizophrenics on 35 major aspects of early history and adjustment. Of these 35 variables, 13 (or 37%) failed to reveal a reliable difference between the two samples. On 5 of the 22 variables which yielded reliable differences, the normals were characterized by greater frequency of the undesirable or pathogenic factor. Specifically, the normals had a greater frequency of poverty and invalidism in their childhood homes, poorer heterosexual adjustment and adequacy of sexual outlet, and a greater incidence of an intellectualized, ritualized orientation toward religion. Additionally, the greater frequency of divorce in the childhood homes of the normals approached reliability.

The schizophrenics were characterized by reliably higher incidence of unfavorable relationships with mothers and fathers, poorer attitudes toward and achievement in school, less occupational success and satisfaction, higher rates of social withdrawal, lack of social adeptness and poise, narrow interests, limited aspiration, vague life plans, and lack of initiative. These personal history characteristics which are predominant in the schizophrenics are in line with the general description which has been made of the preschizophrenic personality and lend some support to the central concept of withdrawal. However, the extent to which these same characteristics were found in closely approximate proportions in the histories of the normals suggests the need for great reservation in interpreting the isolated schizophrenogenic potency of such factors as the mother-child relationship.

The notion that any single circumstance, deprivation, or trauma contributes uniformly and inevitably to the etiology of schizophrenia is called into serious question. The necessity of studying the incidence of such factors in appropriate samples is exemplified. It is suggested that the patterning of life experiences may be more crucial than occurrence or absence of specific psychic stresses. The finding of "traumatic" histories is nearly a fourth of the normal subjects suggests the operation of "suppressor" experiences or psychological processes of immunization. It is suggested that improved insights into mental illness may be afforded by careful, intensive studies of the life histories of normals.

REFERENCES

ALDRICH, C. K., & COFFIN, M. Clinical studies of psychoses in the Navy. I. Prediction values of social histories and the Harrower-Erickson Test. *J. nerv. ment. Dis.*, 1948, **108**, 36–44.

BREUER, J., & FREUD, S. *Studies on hysteria*. J. Strachey (Trans.) New York: Basic Books, 1957.

BROCKWAY, A. L., GLESER, G., WINAKUR, G., & ULETT, G. A. The use of a control population in neuropsychiatric research (psychiatric, psychological, and EEG evaluation of a heterogeneous sample). *Amer. J. Psychiat.*, 1954, **3**, 248–262.

DAHLSTROM, W. G. An exploration of mental status syndromes by factor analytic techniques. Unpublished doctoral dissertation, Univer. of Minnesota, 1949.

LANE, ROBERT C. Familial attitudes of paranoid schizophrenic and normal individuals of different socioeconomic levels. Unpublished doctoral dissertation, New York Univ., 1955.

NIELSEN, G. K. The chlidhood of schizophrenics. *Acta psychiat. neurol. Kbh.*, 1954, **29**, 281–290.

SHAFFER, L. F., & SHOBEN, E. J., JR. *The psychology of adjustment*. (2nd ed.) New York: Houghton Mifflin, 1956.

SCHOFIELD, W., HATHAWAY, S. R., HASTINGS, D. W., & BELL, DOROTHY. Prognostic factors in schizophrenia. *J. consult. Psychol.*, 1954, **18**, 155–166.

STEINBERG, D. L., & WITTMAN, M. P. Etiologic factors in the adjustment of men in the Armed Forces. *War Med.*, 1943, **4**, 129–139.

WALKER, H. M., & LEV, J. *Statistical inference*. New York: Holt, 1953.

WOOLLEY, L. F. Experimental factors essential to the development of schizophrenia. In P. Hoch, & J. Zubin (Eds.), *Current problems in psychiatric diagnosis*. New York: Grune & Stratton, 1953.

Received July 9, 1958.

A CIRCUMPLEX MODEL FOR MATERNAL BEHAVIOR

EARL S. SCHAEFER

National Institute of Mental Health, Bethesda, Maryland

THE goal of this research is to develop an ordered, parsimonious nomological network for the domain of maternal behavior. A theoretical basis for this organization of maternal behavior is provided by Cronbach and Meehl's (1955) analysis of construct validity and by Guttman's (1954) circumplex model for the analysis of the interrelationships of a set of concepts. Both articles suggest that if a systematic set of relationships among concepts can be demonstrated, the scientific usefulness of a conceptual scheme is increased. The purpose of this paper is to demonstrate the generality of a social interaction conceptualization of maternal behavior by ordering the intercorrelation matrices of three sets of data on maternal behavior. When ordered both with factor analysis and with Guttman's circumplex model, similar two-dimensional organizations of maternal behavior concepts were found for the three sets of data. Additional confirmation of the two-dimensional organization was found in Baldwin, Kalhorn, and Breese's (1945) data on the Fels Parent Behavior Rating Scales, although Roff's (1949) factor analysis and Lorr and Jenkins' (1953) second-order factor analysis demonstrate that additional dimensions can also be found in the Fels scales.

DATA AND METHOD

Descriptions of the data ordered in this paper can be found in published studies by Schaefer, Bell, and Bayley (in press), Sanford, Adkins, Miller, and Cobb (1943), and Baldwin, Kalhorn, and Breese (1945). These studies have in common the use of concepts that describe molar social and emotional interactions. This level of conceptualization is supported by Symond's (1939) emphasis upon the quality of the parent-child relationship and by Orlansky's (1949) conclusion that the total pattern of a child's experience rather than specific infant care practices is the important factor in personality development. The relatively molar and abstract concepts used have been given behavioral definitions or have unambiguous interpretations.

The primary statistical model used in the organization of these studies was Guttman's radex theory (1954). Radex theory, which Guttman (1954) calls "a new approach to factor analysis," includes two components, the simplex, which is not pertinent to this paper, and the circumplex. Guttman (1954) suggests dividing factor analysis into the earlier approach of *common*-factors and his notion of *order*-factors, although he points out many connections between the two kinds of factor theories. The circumplex order is described as follows: "It is an order which has no beginning and no end, namely, a circular order. A set of variables obeying such a law will be called a *circumplex* to designate a 'circular order of complexity'." In this article, as in Guttman's presentation of the model, the law of circular order or law of neighboring is applied to "the simplest case of the radex which can be completely portrayed by a simple, two-dimensional diagram." The data that radex theory was designed to order are sets of product-moment correlation coefficients. Guttman developed the circumplex model to organize mental test data but has found "some data from the Minnesota Multiphasic Inventory to form an approximate circumplex."

Guttman (1954) points out the need to define a universe of content before attempting to order data. The universe of content of this paper is the social and emotional behavior of a mother toward an individual child. Therefore, variables that only indirectly relate to maternal behavior with the individual child were excluded from the sets of variables to be ordered. Examples of variables that were excluded are behaviors toward persons other than the child, such as Cooperativeness with Examiners, Sociability, and Dependency, and situational stresses such as Financial Stress.

Although Guttman did not attempt to limit the methods by which a circumplex order can be discovered, he explicitly included inspection of the matrix. Guttman's (1954) statement of this point, although made in relation to the simplex, applies equally to the circumplex if "order" is substituted for "hierarchy." "I submit that a hierarchy, if it really exists, may often reveal itself to inspection alone. Not always, to be sure; we shall see reasons why a hierarchy can sometimes be obscured to direct inspection. But surely the converse holds, if inspection reveals a hierarchy, then the hierarchy exists. . . ."

Another method, not reported by Guttman (1954), which was developed in this study to search for a circumplex order was to choose two variables from the universe of content that had high correlations with other variables in the matrix and zero or low correlations with one another. By plotting the correlations of all variables with the two relatively orthogonal variables one could determine, by inspection, if a circular order of the variables was revealed. Use of this method resulted in the ordering of both positively and negatively correlated variables. Guttman had stated that radex theory could allow for negative correlations although he did not discuss them further in his original paper. All circumplex orders reported in this paper were established by plotting correlations of variables with two orthogonal variables and then, by inspection, rearranging the correlation matrix into circumplex order.

TABLE 1

Circumplex Ordered Correlation Matrix, Unrotated Factor Loadings, and Residuals for Zero to Three-Year Maternal Ratings Judged to Belong to Universe
(N = 56)

	Maternal Behaviors	Unrotated Factor Loadings		Circumplex Ordered Matrix and Residuals															
		I	II	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.	Autonomy	-47	67		00	-06	07	-04	-01	-06	05	12	-01	11	03	-17	01	00	02	04	-03
2.	Ignoring	-87	-08	36		-03	08	-05	-08	-09	06	01	-01	-10	-01	02	04	-01	-02	-02	-04
3.	Punitiveness	-76	-49	-03	67		00	00	00	04	00	-02	-06	-06	01	01	-05	04	02	05	05
4.	Perceives Child as Burden	-74	-49	09	76	90		-12	-03	-10	07	07	00	07	02	-07	01	-09	00	04	-06
5.	Strictness	-66	-50	-07	56	75	62		04	07	-07	-05	01	09	00	05	-08	-02	-02	-06	02
6.	Use of Fear to Control	-67	-57	-08	55	79	75	76		13	-06	00	05	01	-07	-03	04	00	04	02	04
7.	Punishment	-61	-53	-13	48	76	61	73	84		01	-12	07	-02	-06	03	-06	12	06	04	10
8.	Irritability	-57	-67	-13	61	76	82	64	70	72		09	03	-06	01	-08	02	04	04	01	01
9.	Anxiety	21	-69	-44	-11	16	25	15	25	12	43		00	-04	09	-09	06	05	03	00	-03
10.	Intrusiveness	29	-73	-64	-20	08	15	18	28	28	35	56		12	-05	-02	05	-03	02	02	02
11.	Concern about Health	34	-53	-41	-36	-06	08	13	08	05	11	40	61		12	-13	-03	-13	-02	04	-08
12.	Achievement Demand	53	-50	-56	-43	-14	-12	-10	-15	-12	05	54	46	56		-10	-09	03	-06	05	03
13.	Excessive Contact	63	-51	-81	-49	-22	-29	-11	-16	-08	-10	39	53	35	48		01	02	00	-02	02
14.	Fostering Dependency	72	-43	-62	-56	-39	-31	-35	-20	-27	-10	51	57	44	51	68		-02	04	-02	-01
15.	Emotional Involvement	76	-23	-51	-65	-43	-54	-41	-38	-22	-24	37	31	25	55	62	63		04	02	03
16.	Expression of Affection	89	15	-30	-80	-73	-73	-68	-65	-56	-57	12	17	20	33	48	62	69		03	02
17.	Equalitarianism	82	43	-06	-76	-78	-78	-81	-78	-69	-75	-13	-05	09	26	28	39	54	82		01
18.	Positive Evaluation	76	44	10	-74	-75	-84	-70	-72	-59	-71	-17	-08	-05	21	28	35	51	77	82	
				Correlations of Other Variables with Circumplex Order																	
	Rejection of Homemaking			21	67	57	74	33	46	37	61	22	-03	-15	-12	-29	-28	-28	-46	-52	-50
	Negative Emotional States			-12	49	43	49	41	43	30	48	12	10	-06	-12	10	04	-28	-45	-55	-54
	Mood Swings			-27	56	57	59	50	67	57	74	50	42	20	23	05	09	-01	-35	-50	-53
	Financial Stress			-02	31	23	31	32	20	30	36	08	09	00	-19	05	-20	-25	-24	-35	-26
	Poor Physical Health			-14	23	33	42	22	24	31	36	04	20	-01	-11	-07	-18	-17	-21	-28	-23
	Dominance			-44	-16	08	17	05	19	11	30	72	57	48	72	27	47	38	14	-02	01
	Self-Abasement			-18	-14	-05	14	00	04	02	25	40	45	42	17	22	29	05	15	04	-04
	Dependency			-08	18	-01	29	-02	03	02	35	27	23	30	21	00	17	01	06	-06	-12
	Over-Conscientiousness			-53	-48	-16	-12	-02	-13	-06	-02	40	47	64	62	55	42	39	35	23	12
	Sociability			-17	-32	-29	-22	-21	-27	-19	-09	17	26	33	45	17	25	39	46	38	36
	Cooperativeness			22	-63	-62	-64	-47	-55	-54	-69	-20	-23	09	05	11	13	34	56	65	60
	Estimated Intelligence			22	-35	-28	-41	-19	-36	-23	-38	-13	-26	04	23	-01	-18	31	24	37	37
	Narcissism			00	07	-01	02	02	-02	01	14	27	20	16	35	-14	03	07	00	01	05

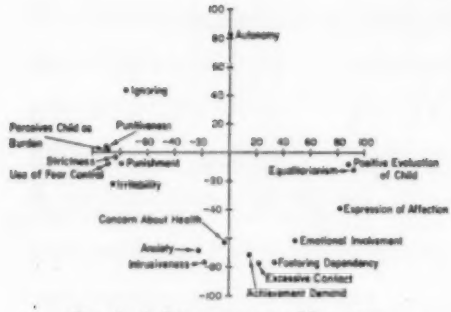

FIG. 1. A CIRCUMPLEX OF MATERNAL BEHAVIOR RATINGS

Rotated factor loadings based on data from observations of maternal behavior between the child's age of zero to three years

To verify the orderings so obtained with a more familiar method, the correlation matrices were factor analyzed using Thurstone's (1947) complete centroid method. In the judgment of the author, the residuals after extraction of two factors do not justify extraction of a third factor since the error variance, due to the small size of the samples, is relatively large. The communalities of practically all variables are relatively high, and the circular order obtained by plotting the loadings of the variables on the two factors is practically identical to the circular order of variables obtained by inspection of the correlation matrix. Although plotting the unrotated factor loadings revealed the circumplex order, orthogonal rotations of two sets of data were made to give the plotted figures a common orientation. These rotations made it possible to interpret the results in terms of both *common*-factors and *order*-factors.

RESULTS AND DISCUSSION

The first circumplex organization of maternal behavior was developed from ratings by Schaefer et al. (in press) of data gathered at the University of California Institute of Child Welfare (Jones & Bayley, 1941). Fifty-six mothers were rated on 32 behavior variables from sets of notes that had been made by an observer during 10 to 20 testing sessions, with each mother and child between the child's age of one month to three years. Reliabilities of the combined scores of three judges ranged from .67 to .89, with a median of .82. Pearson product-moment correlations were computed among the 32 variables. Three judges were asked to select the rating scales that were directly related to the social and emotional behavior of the mother with an individual child. High agreement between judges was found. The 18 scales selected by at least two of the three judges were retained. In order to explore for a circumplex order, the correlations of each variable with Equalitarianism and Intrusiveness were plotted. The variables appeared to fall into a circular order. The correlation matrix was rearranged by inspection into the ordered matrix reported in Table 1. Since inspection revealed, within error limits, a well-ordered matrix, it was concluded that the variables included in this universe form a quasi-circumplex as defined by Guttman (1954). To verify this ordering with a well-known method, the matrix was factor-analyzed with Thurstone's (1947) complete centroid method. The factor loadings and the second factor residuals are reported in Table 1. The second factor residuals are small and do not justify extraction of a third factor. Before plotting the factor loadings in Fig. 1, the factors were orthogonally rotated clockwise 35° 3′. Plots of the factor loadings give a circular order that is identical to that derived by inspection from the correlation matrix.

The findings could be interpreted in terms of two bipolar dimensions which could be labeled *Autonomy* vs. *Control* and *Love* vs. *Hostility*. The first dimension would be represented by Autonomy at one pole and Anxiety of the Mother, Intrusiveness, Concern about Health, Achievement Demand, Excessive Contact, Fostering Dependency, and Emotional Involvement at the other pole. One pole of this dimension resembles the concept of maternal overprotection (Levy, 1943). The positive pole of the second dimension would be Positive Evaluation of the Child, Equalitarianism, and Expression of Affection, and the negative pole would be Ignoring, Punitiveness, Perceives the Child as a Burden, Strictness, Use of Fear to Control, Punishment, and Irritability. The dimension of acceptance vs. rejection is similar to this dimension (Symonds, 1939) although *Love* vs. *Hostility* is suggested to represent it here. However, consideration only of the *common*-factors would not reveal the sequential ordering within dimensions that is shown by Fig. 1.

The correlations with the circumplex ordered variables of the variables that were excluded as not clearly relating to the universe of content are also reported in Table 1. Many of these variables are not highly related to the

TABLE 2

CIRCUMPLEX ORDERED CORRELATION MATRIX, UNROTATED FACTOR LOADINGS, AND RESIDUALS FOR NINE- TO FOURTEEN-YEAR INTERVIEW RATINGS
(N = 34)

Scale No.	Maternal Behaviors	Unrotated Factor Loadings		Circumplex Ordered Correlation Matrix and Residuals																
		I	II	1	15	7	12	18	21	28	8	9	20	13	24	17	5	23	10	27	19
1	Ignoring	-80	44		04	01	05	01	-01	-02	02	06	-03	00	06	03	-02	02	04	00	04
15	Punitiveness	-90	11	01		10	03	-02	-06	01	-03	06	06	04	00	11	01	03	04	-02	03
7	Irritability	-87	07	74	89		10	04	-04	00	-15	05	-05	19	03	17	-09	04	06	02	09
12	Perceives Child as Burden	-42	-03	38	41	46		07	04	03	01	00	-17	04	12	-02	04	-04	05	05	23
18	Use of Fear to Control	-85	-05	67	74	78	43		11	09	-11	-02	-11	-02	04	-01	02	-05	-02	03	09
21	Strictness	-62	-17	42	48	49	30	65		13	-08	00	-24	-17	20	00	05	-14	-11	02	07
28	Wish to Control Child	-82	-46	44	70	68	38	81	72		-08	-03	-05	-08	02	-03	00	-04	-03	-01	01
8	Social Isolation of Child	-67	-38	39	53	40	30	48	40	64		-06	14	-15	00	-13	04	-08	01	-04	-16
9	Intrusiveness	-66	-56	34	59	58	30	57	51	77	59		01	04	05	-06	03	05	05	01	-02
20	Excessive Contact	-41	-64	02	36	27	02	27	12	58	65	64		11	-20	07	01	10	03	-02	-18
13	Fostering Dependency	-16	-60	-13	11	29	13	15	03	33	19	48	55		-08	28	-20	15	01	03	10
24	Achievement Demand	19	-64	-37	-24	-17	06	-09	19	15	11	29	13	27		13	-03	-07	-02	02	05
17	Emotional Involvement	19	-75	-45	-14	-04	-08	-13	-08	15	02	22	47	70	65		-05	09	03	-02	-02
5	Positive Evaluation of Child	65	-55	-78	-63	-69	-21	-50	-26	-28	-19	-09	09	03	44	48		-01	00	-05	-02
23	Expression of Affection	80	-46	-82	-74	-69	-37	-71	-56	-49	-45	-22	06	30	37	58	76		02	00	04
10	Equalitarianism	88	-29	-79	-78	-72	-31	-76	-60	-62	-47	-37	-15	04	33	42	73	85		02	02
27	Positive Mother-Child Relationship	92	-22	-84	-87	-80	-33	-74	-51	-66	-58	-48	-26	01	33	31	67	84	89		13
19	Autonomy	45	62	-05	-30	-26	02	-32	-31	-64	-70	-67	-76	-34	27	-40	-07	12	24	40	
				Correlations of Other Variables with Circumplex Order																	
6	Rejection of Homemaking Role			77	73	78	28	59	24	40	25	31	07	06	-32	-19	-67	-66	-57	-72	-11
14	Anxiety of Mother			44	70	67	33	48	37	65	43	64	42	34	-09	21	-26	-39	-52	-65	-51
26	Emotional Withdrawal of Mother			38	63	56	18	55	43	64	49	51	52	20	-11	20	-18	-43	-48	-58	-52
3	Dependency of Mother			19	36	42	43	14	13	22	16	30	11	42	11	27	-13	-11	-19	-31	-08
4	Sociability of Mother			-50	-66	-51	-18	-52	-25	-66	-58	-46	-58	-02	25	10	34	50	59	67	49
2	Cooperativeness			-76	-88	-80	-43	-71	-47	-75	-57	-67	-46	-14	19	08	48	65	70	83	46
11	Marital Happiness			-55	-73	-73	-17	-52	-29	-50	-38	-38	-44	-23	13	-12	47	54	53	62	30
22	Estimated Intelligence of Mother			-33	-43	-31	-09	-29	-08	-47	-40	-29	-47	-12	43	20	31	33	46	56	44
16	Positive Emotional State of Mother			-31	-36	-38	-30	-36	-14	-37	-22	-35	-13	-01	31	20	22	36	36	45	22
25	Communicativeness of Mother			-40	-58	-37	-36	-42	-34	-58	-56	-50	-44	05	17	08	09	39	39	53	51

circumplex set, while the correlations of others with the circumplex set are relatively high. It seems probable that if variables that do not belong to a universe are included in a matrix they may obscure the circumplex order that exists.

The circumplex organization of maternal behavior was confirmed from a second set of data reported by Schaefer et al. (in press). The data were ratings made from written notes on either one or two home interviews with 34 mothers between the child's age of 9 to 14 years. The subjects were again from the Berkeley Growth Study (Jones & Bayley, 1941). The concepts were revised from the first set of ratings to be appropriate for the age level of the children and for the type of data available. Median reliability after combining the ratings of three judges was .85. The 18 variables selected by at least two of the three judges as belonging to the universe are reported in circumplex order in Table 2. The order was established by plotting the correlations of all variables with two orthogonal variables and then rearranging the correlation matrix by inspection with the plotted order as a preliminary guide. The fact that the ordering of this matrix is not as clear as the first may be due to the fact that the interview data may not be as valid as observation data. Again two factors were extracted from this matrix, and the second factor residuals were computed. Since the small size of the sample suggests that much of the remaining variance may be error variance, a third factor was not extracted. The communalities of most variables are high. The factors were rotated counterclockwise orthogonally 35° 58′ to an orientation similar to that of the first set of data and are plotted in Fig. 2.

The orthogonal rotation did not result in simple structure nor is it obvious that simple structure in terms of *common*-factors exists for this matrix. There is, however, a relatively clear ordering of variables within a two-dimensional space. The rotated reference axes could again be interpreted as *Autonomy* vs. *Control* as defined by Autonomy and the opposed variables of Strictness, Social Isolation of the Child, Wish to Control the Child, Intrusiveness, Excessive Contact, and Fostering Dependency. The horizontal axis might be labeled *Love* vs. *Hostility* as represented by

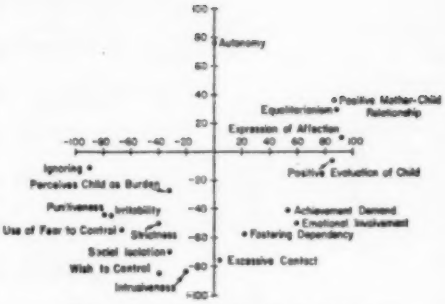

FIG. 2. A CIRCUMPLEX OF MATERNAL BEHAVIOR RATINGS

Rotated factor loadings based on data from interviews with mothers between the child's age of nine to fourteen years

Positive Mother-Child Relationship, Equalitarianism, Expression of Affection, Positive Evaluation of the Child, and Emotional Involvement as opposed to Ignoring, Punitiveness, Irritability, and Use of Fear to Control. However, a summary of these data in terms of *common*-factors seems to the author a less adequate presentation than an organization of the variables in circumplex order, as reported in Fig. 2.

The organization of the interview data on maternal behavior is similar to the organization of the behavior observations. Despite some changes in the intercorrelation of variables, the ordering is quite similar with a sequence of high autonomy, an equalitarian positive relationship, high positive involvement, possessiveness, hostile control, hostile irritability, and hostile ignoring. Although some sectors of the circular space are less well represented than others, by projecting polar opposites of the included variables the entire two dimensional space would be adequately represented.

The ratings by Sanford et al. (1943) of parental press variables were also organized with the circumplex model. The intercorrelations of the variables that were judged to be directly related to the universe are ordered in Table 3. The unrotated factor loadings and second-factor residuals using Thurstone's (1947) centroid method are also reported in Fig. 3 and Table 3. Although the sampling of the universe is less extensive, the ordering obtained can be interpreted as similar to that of the other studies. The area of low Dominance

can be interpreted as autonomy, Deference and Affiliation as a loving positive relationship, Nurturance as a more involved but positive relationship, Dominance and Retention as control, Aggression as hostility, and Rejection and Capricious Discipline as hostile rejection. Presses not included in this universe tend to have lower correlations except for Lack (Family Support) that is highly related to Aggression and Rejection. Some variables of parental behavior are not highly related to the universe as defined. In defining the universe of content as the social and emotional behavior of the mother with the individual child, many variables that are also thought to be relevant to the development of the child have been omitted from these analyses. Although these other dimensions should also be studied intensively, they should not be permitted to obscure the circumplex order that exists within the specified universe.

Published research on the Fels Parent Behavior Rating Scales allows a comparison of different ways of organizing the same set of data. Baldwin et al. (1945) reported the intercorrelations of the scales and also reported an organization of 19 of the variables into three major clusters: Democracy in the home, Acceptance of the child, and Indulgence. They also reported "a number of other syndromes which seem of minor importance." Roff's (1949) factor analysis of the Fels data resulted in seven factors, thus confirming that the total set of data is factorially complex. Lorr and Jenkins (1953) performed a second-order factor analysis using Roff's matrix of correlations of the primary factors and reported three second-order factors that were labeled dependency-encouraging, democracy of child training, and organization and effectiveness of control in the home. Roff's (1949) first two unrotated factors included most of the common factor variance of the scales included in Baldwin et al.'s (1945) three major clusters and also gave a sequential ordering of variables that appeared similar to the circumplex orders of maternal behavior found in previous studies. An attempt was therefore made to organize the set of variables included in Baldwin et al.'s (1945) three major clusters in circumplex order. The quasi-circumplex that emerged is reported in Table 4. The unrotated first-order factor loadings reported by Roff (1949) and the clusters to which the scales were assigned by Baldwin et al. (1945) are reported in circumplex order in Table 5. We see a sequence from the cluster of Indulgence to Acceptance to Democracy. There is no sharp demarcation between these clusters. Most of the variance of the scales is included in the first two factors, and the loadings of the scales on subsequent factors are low. Effective-

TABLE 3

CIRCUMPLEX ORDERED CORRELATION MATRIX, UNROTATED FACTOR LOADINGS, AND RESIDUALS OF PARENTAL PRESS VARIABLES FROM SANFORD ET AL. (1943)

Maternal Behavior	Unrotated Factor Loadings		Circumplex Ordered Correlation Matrix and Residuals							
	I	II	1	2	3	4	5	6	7	8
1. Deference	79	10		01	02	05	05	03	13	−12
2. Affiliation	96	−15	75		−04	−13	−03	07	−05	02
3. Nurturance	75	−50	56	76		07	−03	−12	02	00
4. Retention	−26	−65	−22	−28	19		02	−02	10	−15
5. Dominance	−46	−63	−37	−37	−05	55		06	08	−07
6. Aggression	−96	−15	−75	−83	−76	33	59		−08	16
7. Rejection	−78	38	−45	−86	−75	05	20	61		−11
8. Capricious Discipline	−66	54	−59	−69	−77	−33	−11	71	61	
Correlations of Other Variables with Circumplex Order										
Understanding			00	30	30	05	−12	−41	−26	−33
Exposition			12	14	46	48	56	−22	−12	−24
Excitance			06	37	10	−30	−24	−22	−05	37
Sentience			00	23	11	−10	10	−30	−30	−22
Lack (Family Support)			−45	−59	−75	00	22	71	59	82

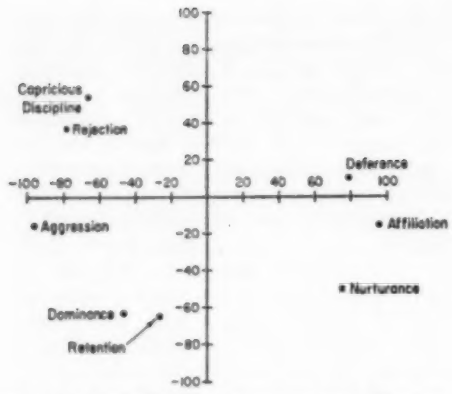

FIG. 3. A CIRCUMPLEX OF PARENTAL PRESS RATINGS Unrotated factor loadings based on data from Sanford, Adkins, Miller, and Cobb (1943)

FIG. 4. A HYPOTHETICAL CIRCUMPLEX OF MATERNAL BEHAVIOR CONCEPTS

ness of Policy and Low Disciplinary Friction have appreciable loadings on the third factor which may describe the success of parental policy with the child.

The three methods of organizing the intercorrelations of this set of data—the circumplex ordering, the cluster method of Baldwin et al. (1945), and Roff's (1949) unrotated first-order factor analysis—give related organizations. These organizations are also related to Roff's (1949) rotated factor loadings and to Lorr and Jenkins' (1953) second-order factor loadings that will not be discussed here. These different ways of organizing the data illustrate the well-known fact that there are many different ways to organize the same set of data. The advantage of the circumplex order over the previous organizations is that it directs attention to the sequential order of variables and does not divide the matrix into discrete clusters or factors.

A generalization of this circumplex model can be used to order other concepts that have been used to describe maternal behavior. A number of these concepts are hypothetically ordered in Fig. 4. Freedom is placed on the *Autonomy* vs. *Control* dimension with its polar opposite, possessiveness. Democratic behavior and cooperative behavior are placed in the quadrant which indicates both loving acceptance and autonomy. Acceptance and its polar opposite, rejection, are placed on the *Love* vs. *Hostility* dimension. Excessive love, overindulgence, protective indulgence, and overprotectiveness are placed in a quadrant that indicates both love for the child and an inability to treat the child as a differentiated individual who has his own activities and interests apart from the parent. The variables of authoritarian, dictatorial treatment and demanding, antagonistic behavior are placed in the quadrant that indicates both a hostile relationship to the child and control; while neglect, indifference, and detachment are placed in the quadrant that indicates varying combinations of hostility and autonomy for the child. This organization seems to satisfy both a law of neighboring and a law of polar opposites.

This paper has demonstrated a two-dimensional organization of maternal behavior concepts which describe molar social and emotional interactions between mother and child. A contrasting approach to the study of parental behavior is that of Sears, Whiting, Nowlis, and Sears (1953), Sears, Maccoby, and Levin (1957), and Whiting and Child (1953), among others, who analyze more molecular variables that relate to the socialization of specific systems of behavior such as oral, anal, sexual, etc. The results of these studies tend to show less organization among variables. Whiting and Child (1953) report practically no relationship between severity of socialization for different systems of behavior in their cross-cultural study. Sears et al. (1953, 1957) also report low

TABLE 4

CIRCUMPLEX ORDERED CORRELATION MATRIX OF THE FELS PARENT BEHAVIOR RATING SCALES FROM BALDWIN ET AL. (1945)

Scale Number	2.11	4.1	1.91	4.2	7.1	2.12	7.2	8.4	8.3	5.2	3.15	3.14	6.1	3.22	8.1	3.16	3.11
2.11																	
4.1	63																
1.91	61	67															
4.2	55	74	71														
7.1	50	53	64	64													
2.12	47	47	63	47	62												
7.2	48	54	69	67	55	51											
8.4	42	47	60	52	46	46	88										
8.3	41	53	60	50	43	53	79	86									
5.2	35	45	51	53	44	49	64	66	70								
3.15	20	24	23	28	33	34	52	50	58	76							
3.14	18	22	27	32	39	39	52	47	54	76	92						
6.1	20	18	29	34	41	37	51	42	44	67	79	87					
3.22	15	18	23	31	37	38	51	48	53	71	88	88	78				
8.1	11	18	20	30	16	27	51	54	61	60	69	71	75	70			
3.16	08	08	26	36	31	29	43	35	36	62	70	74	60	64	60		
3.11	-05	04	08	03	00	10	33	44	49	58	65	64	56	63	56	42	
Variables Reporting Success of Policy																	
3.17	24	30	45	46	30	33	69	70	60	61	49	44	37	43	44	54	34
3.18	21	23	39	45	33	23	66	66	50	54	52	49	41	49	36	57	37

correlations between variables in their studies. Sears et al. (1957) reported a factor analysis of parent behaviors derived from interview data which gave seven factors of parent behaviors, yet the communality of the variables was low. It is interesting to note that these investigators (Sears et al. 1957) report that the quality of maternal warmth is the most pervasive quality of maternal behavior and that it has the most pervasive effects upon the child. Despite the different conceptual schemes, their quality of maternal warmth appears similar to the dimension of *Love* vs. *Hostility* of this study. Child (1954) states, "It is probable that the combined study of general parental characteristics and specific features of socialization as joint antecedent variables will be one of the important directions taken by future research." The tentative integration of molar concepts of parental behavior into an ordered conceptual scheme, presented here, should facilitate such research.

The conceptual model that has been presented may also assist in the interpretation of existing maternal behavior research by suggesting the essential equivalence of diverse conceptual schemes or may help in the selection of variables to be included in future research on maternal behavior. It may challenge investigators to explore new dimensions of

TABLE 5

ROFF'S (1949) UNROTATED FACTOR LOADINGS FOR THE FELS PARENT BEHAVIOR RATING SCALES WITH CLUSTER MEMBERSHIP OF BALDWIN ET AL. (1945)

Scale		Unrotated Factor Loadings							
Number	Scale Name	I	II	III	IV	V	VI	VII	Cluster Membership
2.11	Duration of Contact	68	-20	-06	-07	-06	-15	25	Indulgence
4.1	General Babying	77	-20	-16	-10	-10	02	06	Indulgence
1.91	Child-Centeredness	84	-16	-01	15	10	00	02	Indulgence, Acceptance
4.2	General Protectiveness	80	-06	05	16	-18	03	27	Indulgence
7.1	Solicitousness	72	-01	-17	12	-05	-38	-09	Indulgence
2.12	Intensity of Contact	69	04	-24	22	10	-09	-39	Indulgence
7.2	Acceptance	86	17	23	-05	-03	13	01	Indulgence, Acceptance
8.4	Rapport	80	18	33	-19	07	19	-02	Acceptance
8.3	Affectionateness	79	25	04	-24	15	25	-12	Acceptance
5.2	Approval	69	53	03	00	15	-06	15	Acceptance, Democracy
3.15	Democracy	46	81	00	-12	05	-11	-10	Democracy
3.14	Justification of Policy	47	83	-07	02	-01	-10	-14	Democracy
6.1	Readiness of Explanation	45	78	-14	11	-11	-02	03	Democracy
3.22	Non-Coerciveness	45	79	-05	-04	09	-17	01	Democracy
8.1	Understanding	41	68	-05	-03	-18	35	25	Democracy
3.16	Clarity of Policy	36	68	24	35	-16	-14	09	Democracy
3.11	Non-Restrictiveness	21	67	08	-29	31	19	-16	Democracy
		Variable Reporting Success of Policy							
3.17	Effectiveness of Policy	59	27	53	14	20	15	-10	Acceptance
3.18	Low Disciplinary Friction	53	33	63	-02	-03	-10	-10	Acceptance

maternal behavior rather than to reinvestigate known dimensions with new names.

SUMMARY AND CONCLUSIONS

By organizing several empirical studies of maternal behavior using Guttman's (1954) circumplex model, it was demonstrated that many of the existing concepts of maternal behavior can be ordered within a two-dimensional space. Concepts that do not describe molar social and emotional interactions of mother and child often do not show a clear circumplex ordering. The following interpretations of the findings were made:

1. Two major dimensions of maternal behavior can be isolated in all studies; these can be labeled *Love* vs. *Hostility* and *Autonomy* vs. *Control*.

2. A circumplex organization of this universe shows a clearer nomological network than an organization in terms of discrete dimensions.

3. A theoretical generalization of the circumplex of maternal behavior is presented. It provides an ordered, parsimonious nomological network or model for much of the variance of a mother's social and emotional behavior toward an individual child.

REFERENCES

BALDWIN, A. L., KALHORN, JOAN, & BREESE, FAY HUFFMAN. Patterns of parent behavior. *Psychol. Monogr.*, 1945, **58**, No. 3 (Whole No. 268).

CHILD, I. L. Socialization. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. II. Cambridge, Mass.: Addison-Wesley, 1954.

CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychol. Bull.*, 1955, **52**, 281–302.

GUTTMAN, L. A new approach to factor analysis: The Radex. In P. F. Lazarfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, Ill.: Free Press, 1954.

JONES, H. E., & BAYLEY, NANCY. The Berkeley growth study. *Child Develpm.*, 1941, **12**, 167–173.

LEVY, D. M. *Maternal overprotection*. New York: Columbia Univer. Press, 1943.

LORR, M., & JENKINS, R. L. Three factors in parent behavior. *J. consult. Psychol.*, 1953, **17**, 306–308.

ORLANSKY, H. Infant care and personality. *Psychol. Bull.*, 1949, **46**, 1–48.

ROFF, M. A factorial study of the Fels Parent Behavior Scales. *Child Develpm.*, 1949, **20**, 29–45.

SANFORD, R. N., ADKINS, MARGARET M., MILLER, R. B., & COBB, ELIZABETH. Physique, personality, and scholarship. *Monogr. Soc. Res. Child Develpm.*, 1943, **8**, No. 1.

SCHAEFER, E. S., BELL, R. Q., & BAYLEY, NANCY. Development of a maternal behavior research instrument. *J. genet. Psychol.*, in press.

SEARS, R. R., MACCOBY, E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.

SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child rearing antecedents of aggression and dependency in young children. *Genet. psychol. Monogr.*, 1953, **47**, 134–234.

SYMONDS, P. M. *The psychology of parent-child relationships*. New York: Appleton-Century, 1939.

THURSTONE, L. L. *Multiple-factor analysis*. Chicago: Univer. Chicago Press, 1947.

WHITING, J. W. M., & CHILD, I. L. *Child training and personality: A cross-cultural study*. New Haven: Yale Univer. Press, 1953.

Received July 14, 1958.

PERSONALITY DETERMINANTS OF REPRESSIVE AND OBSESSIVE REACTIONS TO FAILURE-STRESS[1]

ALBERT J. CARON AND MICHAEL A. WALLACH

National Institute of Mental Health and Massachusetts Institute of Technology

MUCH research effort has recently focused on the personality correlates of certain cognitive responses obtained under conditions of ego threat and assumed to reflect differences in defensive predispositions. One form of threat, the exposure of intellectual incompetence, has received special attention in this regard, with defensive reaction being inferred either from incidental recall scores, or perceptual recognition thresholds, for success and failure material. A preference for success items, in recall or perception, has tended to be regarded as indicative of a *repressive* reaction to failure, and preference for failure items, as reflecting either an adaptive coping with failure or an *intellectualization* defense (i.e., preoccupation with the painful event in an attempt to reinterpret its implications for the self).

It is not our purpose here to question these interpretations of the obtained cognitive patterns (this has been done elsewhere (Caron & Wallach, 1957) but, rather, to determine the particular personality dimensions which are responsible for them. While various personality measures have been shown to relate significantly to recall and perceptual preferences secured under failure-stress, a coherent dynamic picture is yet to emerge from the total pattern of these relationships. For one thing, indices of manifestly different, if not opposite, personality dimensions (e.g., hysteria [Eriksen, 1954] and ego strength [Alper, 1948, 1957; Eriksen, 1954]) have been shown to correlate identically with defensive indicators; and, conversely, indices of apparently related variables (e.g., ego-strength and achievement motivation [Atkinson, 1953]) have yielded opposite correlations. The situation is further complicated by the fact that investigators have in some cases used very different instruments to measure the same, or similar, personality constructs; and, in others, somewhat similar instruments for different constructs. Were the extent of dimensional overlap between these various measures known, much of the apparent confusion would undoubtedly vanish. But at this point only meager empirical evidence exists.

The present study attempts to provide some of this evidence by factor-analyzing a broad spectrum of indices from three personality areas which have received the major attention in previous research: the areas of *ego strength*, *need Achievement* and *hysteria*. To determine their significance for defense, the factors that emerged from this analysis were correlated with selective recall scores for a highly stressful interrupted-task situation and with perceptual recognition thresholds for material related to this situation.

METHOD

Subjects

Eighty-four Harvard freshmen, randomly selected as described in our previous paper, received the complete battery of tests used in the factor analysis. Since only half of these *Ss* were still under failure-stress during the recall and perceptual procedures—the other half having been relieved of stress for the purposes of the previous study—correlations between the factorial scores and the cognitive measures are based only on the former 42 *Ss*. Factorial scores were also computed for an additional 33 randomly selected freshmen, whose test scores were not included in the factor analysis but whose cognitive responses were obtained under neutral experimental conditions.

Personality Indices

Ego-strength measures. Alper's ego-strength syndrome (Alper, 1957), adapted from Murray (1942), served as a general guide for the selection of tests in this area, most of the pertinent indices consisting of various subscales of the Edwards Personal Preference Schedule (Edwards, 1954) and the Guilford Inventory

[1] This study was conducted while the senior author held a USPHS postdoctoral research fellowship at Harvard University. He is indebted to Ray Hyman and Charles McArthur for their advice and counsel concerning many aspects of this endeavor and to the Harvard University Health Services for subsidizing the factor analysis reported herein. Grateful acknowledgement is also made to Arthur Couch for facilitating machine computation of these data.

of Factors STDCR (Guilford, 1940).[2] Thus, for ego strength proper ("the power to persist . . . an unconquerable will"), the need Endurance scale of the Edwards Schedule was considered most relevant (Index 12 on Table 1). For need Recognition and need Dominance we employed the need Exhibition (Index 13) and need Dominance (Index 15) scales respectively of the Edwards. With regard to Dejection Pessimism ("the tendency to worry about shortcomings, to be pessimistic about life, to recover slowly from disappointment"), a number of measures in the general domain of neuroticism and anxiety were selected: first, the Guilford Factors STDCR (S measuring social shyness (Index 22), T, introspectiveness and daydreaming (Index 23), D, depressive tendencies (Index 24), C, emotional instability (Index 25), and R, carefreeness—the opposite of dejection pessimism (Index 26)); second, the Psychasthenia scale of the Minnesota Multiphasic Personality inventory (Hathaway & McKinley, 1943), which reflects obsessive self-concern and worry (Index 27); third, the Knutson Personal Security Inventory (Knutson, 1952), which provides an index of security in a number of basic areas (Index 30); and fourth, impairment of digit-retention after failure-stress (scored as the pre- to poststress difference in digit retention, as described in our earlier paper)—assumed to reflect the ability to recover from disappointment (Index 29). The last component of Alper's syndrome represented in the present analysis is Counteractive Achievement, i.e., the tendency to "set difficult goals for oneself and attempt to reach them," "to overcome obstacles . . .," and "to redouble one's efforts in the face of frustration." To measure this variable, a level of aspiration score was constructed (Index 6), consisting essentially of the relative increase in level of aspiration following a performance drop induced by false norms (see below for complete description).

Finally, to determine the relationship between Eriksen's Rorschach index of ego strength (Eriksen, 1954) and Alper's indices, the Rorschach procedure (Index 31) was also administered.[3]

Need Achievement measures. McClelland, Atkinson, Clark, and Lowell (1953) have defined achievement motivation as "concern with success in competition with some standard of excellence" (p. 110). The present study employed a number of indices of such motivation. One was the need Achievement scale of the Edwards Schedule (Index 10). Another was a measure of academic achievement (scored as the difference between attained and predicted rank standing in the Harvard freshman class), which McClelland et al. (p. 237) report is correlated with their TAT index of need Achievement (Index 2). A performance measure also found to correlate with this index (p. 227) and used here, was output during the middle, "sag" period (5th to 7th min.) of an anagrams task (constructing words from the master word, GENERATION), corrected for regression on output during the first 3 min. (Index 8). Another score derived from this anagrams procedure, and assumed to reflect need Achievement, was total output, corrected for regression on the verbal portion of the Scholastic Aptitude Test (Index 7).[4]

Still another measure assumed to be relevant to need Achievement was the over-all goal discrepancy score (Index 5) in the level-of-aspiration procedure mentioned above. Further, following the lead of the McClelland and Liberman (1949) study, perceptual indices too were employed; namely, thresholds for success (Index 4) and failure (Index 3) words (compared with neutral words—see description below). As a final indicator of need Achievement, the dichotomous variable of public vs. private high school attendance (Index 1) was used. Both rational and empirical considerations suggest that Harvard public school boys are highly motivated to achieve. They have consistently stood at or near the top of their high school classes, and emerge from a middle-class cultural matrix wherein the accomplishments of the new generation are expected to surpass those of the old. Harvard private school boys, on the other hand, come from families that have already achieved success, where what a person *is* takes precedence over what he may *become* (cf., McArthur, 1954).

Hysteria-hypnotizability measures. The hysteria scale of the MMPI (Index 28) provided the most direct measure in this area. In line with White's report on various personality correlates of hypnotizability (White, 1938), several subscales of the Edwards Schedule may be regarded as providing indirect measures of hypnotizability. White suggests that there are two types of trance states, the "alert" and the "sluggish." Related to the former is the dimension of deference-autonomy, with susceptible *S*s being high in such Murray needs as Deference (Index 19), Affiliation (Index 17), and Nurturance (Index 16), and low in Autonomy (Index 21) and Aggression (Index 14). The dimension of passivity-activity is related to the sluggish trance, with positive correlations for such variables as n Succorrance (Index 20) and n Abasement (Index 18), and negative for n Order (Index 11). All of these scales appear in the Edwards Inventory and were included in the present analysis. In addition, as another possible

[2] Since various practical considerations precluded administering both the Edwards Schedule and Alper's Psychological Insight Test (Alper, 1957), the former was selected because (*a*) it controlled for social desirability of responses, (*b*) it covered areas deemed relevant to need Achievement and hysteria as well as ego strength, and (*c*) in her validating study Alper used clinical ratings as a superseding criterion of ego strength.

[3] Eriksen's index is based on the McReynolds Concept Evaluation Technique (McReynolds, 1951), ego strength being defined as the extent of agreement between *S*'s judgments of the "reasonableness" of particular Rorschach responses and norms of such judgments. Eriksen's norms were used in the present study.

[4] The anagrams procedure was repeated about three weeks later, using AUTOMOBILE as the master word. Fairly good reliability was obtained for the total output measure, Pearson's r being .78. The corresponding correlation for the "*sag*" scores, however, was only .15. Since any number of factors (e.g., anticipation of less time allowance, conscious efforts to better one's previous score, etc.) could have seriously altered the second sag score (and yet not have affected total output), we decided to include both sag scores (Indices 8 and 9) in the factor analysis.

indicator of deference-autonomy, the short form of the F scale (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950) was administered (Index 32).

Intelligence Measures

To determine whether defensive reactions to the threat of intellectual incompetence might be related to actual intelligence, and whether, in addition, our personality factors were independent of intelligence, we included a number of intellectual measures in the present factor analysis. These were the mathematical (MAT) and verbal (SAT) components of the Scholastic Aptitude Test (Indices 35 and 34 respectively), the Thurstone version of the Gottschaldt Concealed Figures Test (Index 36), Thurstone's Inductive Reasoning Test (Index 37), and finally the prestress scores of the Wechsler digit retention test (Index 38).

Response Stereotypy

One index, not readily rationalizable as relating to ego strength, need Achievement, Hypnotizability, or Intelligence (but sharing some variance with the battery), was also included in the present analysis. This was a "response stereotypy" score developed by Hyman and Jenkin (1956), which measures the tendency to repeat one's previous responses in a two-choice guessing situation (Index 33). Such a measure might conceivably reflect a tendency toward independence and autonomy, but we defer judgment on this point.

Behavioral Indices

Level-of-aspiration procedure. The level-of-aspiration procedure was introduced as a test of verbal fluency. Since it was deemed desirable to keep failure-stress at a minimum prior to the actual stress experience, it was emphasized that fluency had little to do with general intellectual ability. *S*'s task was to list as many different four-letter English words as possible (all beginning with a particular letter), in 1¼ min. After the first such attempt, *E* counted the number of words *S* had produced, then pretended to scan a bogus table (containing "percentile norms for Harvard freshmen") for *S*'s score, recorded it, and in an offhand way asked *S* if he'd like to know how he had done. No matter what *S* responded, *E* read the percentile score of 59 and also explained the meaning of such a score. *E* made some reassuring remarks about *S*'s performance (explaining that he was very likely closer to the 90th percentile relative to *all* college freshmen) and then asked, "Just for the fun of it, how do you think you'll do on the next letter, that is, what percentile score do you think you'll hit?" This procedure was repeated five more times, all *S*s receiving the following sequence of bogus scores: 59, 64, 68, 62, 69, 74. The sequence of letters was arranged in such a manner, and the actual time so manipulated, to insure that *S*'s subjective experience would approximate the bogus levels in each instance. All *S*s were tested individually. Two scores are used in the present analysis. These are (*a*) the total goal discrepancy score (Index 5), i.e., the sum of the goal discrepancies for all six trials; and (*b*) the counteractive achievement score (Index 6) mentioned above in connection with ego strength, which is the goal discrepancy score after the drop to the 62nd percentile (Trial 4), corrected for the regression of this score on the goal discrepancy score for Trial 2 (64th percentile).

Perceptual thresholds for success and failure words. Perceptual recognition thresholds were determined for success- and failure-related words and for neutral words of matched structure and frequency (according to the Thorndike-Lorge L count). The critical words and their matched controls (in parentheses) are as follows: *success words:* succeed (cottage), brilliant (treatment), superior (relative), excellent (excitement); *failure words:* failure (flavor), stupid (signal), unable (useful), defeat (formal). The *S*s, tested individually, were seated 5 ft. from a ground-glass screen in a room illuminated solely by a 40-watt fluorescent ceiling bulb 1 ft. behind *S*. Each word was flashed at $\frac{1}{100}$ sec., beginning at subthreshold brightness levels and rising in two-volt steps until correct recognition was achieved. Half the success and half the failure words, together with their neutral counterparts, were administered in a counterbalanced order in each of two sessions. A raw score for each critical word was first determined as the threshold difference between the critical and matched neutral word (each measured in terms of number of steps required for correct recognition). To prevent any single critical word from having an undue influence on the final threshold score, the median threshold difference was then determined for each critical word, and a deviation of plus or minus 6 from this median was the maximum allowed in computing final scores. The sum of the threshold difference scores of the four success words constituted *S*'s threshold score for success words (Index 4), and the corresponding sum for the four failure words constituted his failure threshold score (Index 3).

Cognitive Tests

Scrambled sentences. Our 84 *S*s were given a bogus intelligence test consisting of 16 scrambled sentences, only half of which were solvable. A detailed description of the test, the method of inducing stress, and the recall procedure, may be found in our earlier report. Thirty-three other *S*s participated in a neutral, task-oriented version of this procedure, which is also described in the previous report. The recall measure most pertinent for the present study is the recall difference score, which consists of the number of completed minus the number of incompleted tasks recalled (CR-IR).[5] As mentioned above, half the stress *S*s were relieved of stress just prior to recall and, hence, are not included in the correlation with our factorial scores.

Perceptual thresholds for test-related words. Immediately after recall, recognition thresholds were determined (for the stress and neutral *S*s) for words pertaining in a general way to the scrambled sentence test, and for neutral words of matched structure and frequency (according to the Thorndike-Lorge L count). The former group included the following words: ***sentence, booklet, phrases, examination, reasoning, scrambled,***

[5] This score was appropriately corrected for *S*s who failed to solve all the solvable tasks. There was no *S* who solved fewer than 6 of the 8 solvable tasks.

and *tested*. The corresponding neutral words were: *sunshine, cookies, sketches, agriculture, leisurely, channeled,* and *regard*. For further details concerning this procedure, our previous paper may be consulted.

Sequence of Procedures

All *S*s participated in seven experimental sessions, the first two of which were individually administered. The first five sessions were administered by the senior author, although the *S*s were told that the personality questionnaires presented in the fourth and fifth sessions were for an unrelated study being conducted by another investigator. The last two sessions were actually administered by this investigator as part of his research.

The sequence of procedures in each session was as follows:

Session I (1 hr.)
1. Judgment of Rorschach responses; 2. First anagrams task; 3. Thresholds for success and failure words (first half)

Session II (1 hr.)
4. Second anagrams task; 5. Thresholds for success and failure words (second half); 6. Level-of-aspiration procedure

Session III (2½ hrs.)
7. Rote memory for digits (prestress); 8. Scrambled sentence test; 9. Rote memory for digits (poststress); 10. Incidental recall of scrambled sentences; 11. Thresholds for exam-related words

Session IV (1¼ hrs.)
12. F scale; 13. Edwards Schedule

Session V (¾ hr.)
14. Hysteria and Psychasthenia scales of the MMPI

Session VI (1½ hrs.)
15. Guilford Inventory; 16. Knutson Inventory; 17. Gottschaldt Hidden Figures Test; 18. Thurstone Inductive Reasoning Test

Session VII (1 hr.)
19. The two-choice guessing situation from which Hyman's Response Stereotypy score is derived.

RESULTS

Factor Analysis of Personality Data

The scores of all tests were converted to single-digit (0–9) coded scores. Each *S*'s coded scores for all 38 variables were entered on an IBM card, and the intercorrelations between variables (Pearson product-moment coefficients) were computed by standard machine techniques. Factoring was done by Thurstone's centroid method, and the criteria of rotation were simple structure and psychological meaningfulness.[6]

[6] Both the factorial and rotational computations were machine processed according to programming techniques recently developed by Arthur Couch.

TABLE 1
FINAL ROTATED FACTOR LOADINGS[a]

Measure	I	II	III	IV	V	VI	h^2
1. High school	02	−10	−25	−01	−37	−11	22
2. Ac. ach.	01	18	09	−21	41	06	26
3. Failure limen	−09	−04	−18	23	26	−24	22
4. Success limen	04	10	−18	03	23	−54	39
5. Total L.A.	11	−03	26	−20	36	−25	31
6. L.A. 4/2	09	19	09	17	41	−24	31
7. Total anag.	01	73	14	05	10	−02	57
8. Anag. sag$_1$	10	42	08	−11	16	09	24
9. Anag. sag$_2$	−14	47	09	01	12	00	26
10. n Ach.	26	09	−03	−50	26	12	41
11. n Ord.	−15	27	−11	−13	07	−39	28
12. n End.	04	−05	−08	01	66	01	45
13. n Exh.	−18	−03	47	07	−26	−14	35
14. n Aggr.	31	09	16	−18	−43	−19	38
15. n Dom.	−22	04	13	−29	−23	−18	24
16. n Nur.	−07	−07	−07	76	07	18	63
17. n Affil.	−20	−05	−02	42	20	05	26
18. n Abas.	14	−08	−19	38	09	16	24
19. n Def.	−06	−10	−20	24	25	12	19
20. n Succor.	08	09	−23	08	−20	−16	14
21. n Auton.	09	08	16	−33	03	10	16
22. S	45	11	−67	−01	−02	−07	67
23. T	57	−27	−04	−27	14	17	52
24. D	91	−14	−12	00	08	−06	87
25. C	89	−12	07	03	06	−10	83
26. R	−04	12	56	13	−16	00	37
27. Psych.	82	−16	−18	−03	−06	−08	74
28. Hyst.	25	−28	12	18	−03	29	27
29. Dig. imp.	03	−10	18	26	−15	02	13
30. Knut.	−43	00	29	−03	05	36	40
31. Rorsch.	−10	20	18	−07	−06	15	11
32. F	−12	14	−03	20	−03	−40	24
33. Stereo.	−19	06	33	07	−04	13	17
34. SAT	25	15	−36	−37	16	14	40
35. MAT	−05	58	−26	−08	14	−07	44
36. Conc. fig.	−17	60	−30	−06	03	12	50
37. Reas.	02	69	−06	−04	02	01	48
38. Dig. ret.	10	37	−19	14	08	05	21

[a] Decimal points are omitted.

Table 1 presents the final rotated factor matrix for the 38 variables.[7] Six factors emerged from the analysis, of which five appear to be significant.

Factors I and III are familiar and offer no real interpretation problem. Factor I is the dimension labeled *neuroticism* by Eysenck (1953). Guilford's S, T, D, and C factors as well as the Psychasthenia scale of the MMPI define its maladjustment pole, and the Knut-

[7] The correlation matrix and the centroid loadings have been deposited with the American Documentation Institute. Order Document No. 5964 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

son Inventory, its adjustment pole. This factor also seems related to the "surgency-desurgency" polarity described by Cattell (1956). Factor III is the *extraversion-introversion* dimension, also identified by Eysenck. It is positively anchored by Guilford's R and the need Exhibition scale of the Edwards, and loads negatively in Guilford's S. This factor may be related to Cattell's Source Trait II (immunity vs. reactivity to threat), which emphasizes the carefree, limelight-seeking aspect of extraversion.

Factor II obviously concerns intellectual ability. Its high-loading tests are total word output on the anagrams task (SAT partialled out), the Thurstone Inductive Reasoning Test, the Gottschaldt Hidden Figures Test, and the Mathematical Aptitude Test. Both the anagrams and the Gottschaldt tests involve the ability to restructure an initial organization into new forms, thereby suggesting that the particular intellectual dimension involved here is *cognitive flexibility*. This capacity would also appear to underlie performance on the Inductive Reasoning Test, where one must superimpose a sequential structure on a series of events, and undoubtedly much the same process enters into the solution of mathematical problems. It is notable also that the McClelland-Clark variable of anagram output during the middle "sag" period received its highest loading in this factor (for both anagram tasks), and thus seems to be more strongly influenced by intellectual flexibility than by achievement motivation.

Our fourth factor is heavily saturated with variables from the Edwards Schedule—needs Nurturance, Affiliation, and Abasement defining the positive pole, and need Achievement and, to some extent, need Autonomy, the negative pole. This dimension might be termed *altruism vs. narcism*, or, more broadly, *other-orientation vs. self-orientation*, the positive pole representing concern for others and depreciation of self, the negative pole, autonomous self-seeking. At another level, one may distinguish between the need for togetherness and cohesion with others, and the need for distinction and respect *from* others.

Factor V loads positively in the Edwards need Endurance scale, Academic Achievement, and Counteractive Achievement (as measured by our level-of-aspiration score); it loads negatively in the Edwards needs Aggression scale and Private School attendance.[8] The common theme here seems to be persistent striving for achievement vs. minimal achievement effort and readiness to disparage the accomplishments of others. The high striver would seem to bolster self-esteem by outstripping others, the low striver, by reducing others to a level lower than himself. In many respects this dimension approaches the Alper-Murray concept of ego strength ("... the power to persist ... the tendency to overcome obstacles ... a successful need Achievement ..."). At the same time, however, it contains many of the indices of need Achievement. In this light we will use the more descriptive term, *perseverance for achievement*, for this dimension.

In sum, then, the five factors isolated in our analysis may be tentatively labelled *Neuroticism*, *Intellectual Flexibility*, *Extraversion*, *Other-orientation*, and *Perseverance for Achievement*. Consider next the relation between these factors and our cognitive indicators of avoidant and obsessive defensive reactions.

Correlation of Factorial Scores with Recall and Perception

Factorial scores for each *S* were computed in terms of those tests which loaded almost exclusively in one factor and at a value of .35 or higher. If a test loaded heavily in one factor and moderately in another (e.g., Guilford's S), it was used only for the factorial score of the former. Thus, the tests selected for each factor were as follows: *Neuroticism*—Guilford's T, C, and D factors, and the MMPI Psychasthenia scale; *Intellectual Flexibility*—total anagram output, Thurstone's Inductive Reasoning Test, the Gottschaldt Hidden Figures Test, and the MAT;[9] *Extraversion*—Guildord's R and S, and the need Exhibition scale of the Edwards Schedule; *Other-orientation*—the Edwards need Nurturance, Affiliation, Abasement, and

[8] In our earlier report we indicated that Public-Private school was the dominant variable in a rotated factor. Subsequent reanalysis suggested the advisability of a new rotation which led to the factor reported above. This of course has no bearing on the functional relationships demonstrated in the previous report.

[9] The two middle output scores of the anagrams test were not used, since they apparently reflect the same process that contributes to total output. The digit retention test was not included, since we already had four tests with strong loadings.

TABLE 2

CORRELATIONS BETWEEN FACTORIAL SCORES AND COGNITIVE MEASURES FOR SUBJECTS UNDER FAILURE STRESS

Factor	Cognitive Measure		
	PT	CR-IR	TR
I	−.10	−.04	.17
II	.01	−.05	−.05
III	.23	.02	.08
IV	−.01	.20	.06
V	.47	−.10	.07

Achievement scales; *Perseverance for Achievement*—the Edwards need Endurance and need Aggression scales, the academic achievement score, the level-of-aspiration counteractive achievement score, and private school attendance. The factorial score for each *S* was the algebraic sum of his coded scores on the tests pertinent to each factor.[10]

Table 2 presents the correlations between the factorial scores of each of our five factors and the recall direction (CR-IR), total recall (TR), and average perceptual threshold (PT) scores of the 42 *S*s who were still under failure-stress when the cognitive measures were obtained. These results may be quickly summarized. With regard to PT, the correlation of .47 with Factor V, Perseverance for Achievement, is significant well beyond the .01 level for a two-tail test, persons high on Perseverance showing lower recognition limens for material reminiscent of the test situation. No other factor is significantly correlated with PT, although Factor III, Extraversion, approaches significance with an r of .23.[11] Persons high in Extraversion, then, may also tend toward lower recognition limens for test-related material. The multiple correlation coefficient between Factors III and V and the perception measure is .52.

[10] The original coded scores were not transformed to standard scores, since the standard deviations for all tests were approximately equal. To equate the contribution of the private-public school dichotomy to the perseverance factorial score with the other tests in this factor, that dichotomy's coded scores of 1 and 0 were transformed to 4 and 0 (the differential of 4 approximating the average interquartile range of the other tests).

[11] With our sample size and a two-tail test, the .05 and .01 significance levels require an r of .30 and .39 respectively.

TABLE 3

COGNITIVE MEANS FOR HIGH VS. LOW SCORERS ON THE PERSEVERANCE FACTOR UNDER NEUTRAL AND FAILURE-STRESS CONDITIONS

	PT		CR-IR	
	Neutral	Stress	Neutral	Stress
High perseverance scorers	2.4	−1.4	.66	.24
Low perseverance scorers	0.4	4.1	.42	.98

None of the factors is significantly correlated with CR-IR, and the only one approaching significance is Factor IV, Other-orientation, with an r of .20 (other-oriented persons tending to recall more successes than failures). With regard to total recall, only Neuroticism shows any trend (r = .17), neurotic *S*s tending to yield higher total recall under stress conditions than nonneurotic *S*s.

The correlation between our perception and recall measures is −.23, which is in the expected direction and approaching significance. It indicates some tendency for persons with lower recognition limens for test-related content to recall test failures better than successes.

Mean scores on the cognitive measures for high and low Perseverance *S*s in both stress and neutral conditions are shown in Table 3. The cutting point for high vs. low scores was the median perseverance score for the *S*s in stress, relief, and neutral groups combined. For PT, the interaction between conditions and Perseverance score is significant, perceptual recognition limens for test-related material increasing from neutral to stress conditions for low Perseverance *S*s, but decreasing from neutral to stress conditions for high Perseverance *S*s. (Snedecor's Method of Unweighted Means [Snedecor, 1946, p. 287] yields an interaction F of 4.11, while 3.98 is required for the .05 level of significance with 1 and 70 *df*.) In addition, the difference between limen means for high vs. low Perseverance *S*s under the stress condition is significant beyond the .05 level by a two-tail t test. The comparable interaction for CR-IR is not significant, although the trends of the means are in the appropriate directions. Further, the analogous difference between recall means for high vs. low Per-

severance *S*s under stress approaches significance by a *t* test, the two-tail *P* being less than .25.

Cognitive means were also computed for high and low subgroups on the remaining four factors, for both neutral and stress conditions. The only near-significant difference occurs for the perceptual recognition limens of high vs. low scorers on the Other-orientation factor under neutral conditions. Low scorers tend toward higher limens than high scorers (the means are −0.5 and 3.8 respectively), the two-tail *P* level being just above .10 by a Mann-Whitney *U* test.

DISCUSSION

Hysteria, need Achievement and Ego Strength

None of our factors seems particularly related to Hysteria as traditionally conceived. The Hysteria scale of the MMPI loaded equally in Neuroticism and Intellectual *In*flexibility, but otherwise shared little variance with the test battery. White's suggestion of a passivity-activity dimension and a deference-autonomy polarity—each related to different aspects of hypnotizability—was not supported, although again much of the variance of the Edwards Succorance, Deference, and Autonomy scales was unique. Thus, while the dimensionality of these scales and the Hysteria scale is still unclear, our results do indicate that these indices have very little to do with each other.

With respect to need Achievement, most of the performance variables assumed to be pertinent to this area defined the positive pole of the Perseverance for Achievement factor (along with the Edwards need Endurance scale). That the Edwards need Achievement scale was mostly associated with the self-oriented pole of Factor IV, and loaded negligibly on Perseverance for Achievement, may result from the fact that this scale measures *value* Achievement rather than *need* Achievement—a suggestion made by McClelland.[12] Perceptual sensitivities to success and failure words (obtained, as the reader will recall, under neutral conditions) were only minimally related to Perseverance for Achievement. This finding, coupled with the fact that the perseverance factor did relate strongly to perceptual thresholds for concepts referring to a failure experience, suggests that the processes determining such sensitivities are operative only under conditions of motivational arousal (a point also recently stressed by Atkinson, 1958). The loading of public school attendance at the positive end, and private school attendance at the negative end, of the Perseverance for Achievement factor parallels the picture of the academic pattern of these two Harvard groups obtained via projective techniques by McArthur (1954).

The preceding discussion may have implied that we regard the factor labeled Perseverance for Achievement as synonymous with need Achievement, and certainly it is reasonable to assume that individuals scoring high on this factor are also high in achievement need. The converse, however, is not necessarily the case; that is, among those persons who are "concerned with success" (McClelland's definition of need Achievement) may be some who for various reasons do not "strive" for success in the face of obstacles. One reason for the failure to counteract obstacles may be a lack of sufficient self-discipline or persistence.

The concept of ego strength, as well as its historical antecedent, the concept of will power, would seem to involve precisely this element of persistence vis-a-vis goals. As such, it perhaps may best be conceptualized as a "style," in the sense that Klein (1951) uses the term—in this case, a style of coping with respect to goals and obstacles. The perseverance factor isolated in the present analysis would appear to be a composite of such a strong ego style plus orientation toward the particular goal area of success. This may result from the fact that when American investigators construct tests to measure persistence and endurance, they typically concretize them in terms of the valued goals of their own culture.

Personality and Defensive Orientation

We have seen that one of our personality factors—Perseverance for Achievement—bears a strong relationship to defensive orientation, Lows reacting to stress repressively and Highs reacting vigilantly. This relationship exists only with regard to our perceptual index and, while the direction with respect to recall is consistent, it is not significant. Several matters

[12] Personal communication.

require discussion at this point: 1. Why was there no equally strong relationship with the recall measure? 2. How do our findings bear on related studies? 3. What particular processes underlie the relationship between Perseverance for Achivement and type of defense?

1. That perseverance differences are not significantly related to selective recall of specific test items becomes understandable when one considers the findings reported in our previous paper. It was shown there that a true repression effect occurred for the test as a total experience rather than for particular items (item preference being a function of registration rather than mnemonic selectivity). Indeed it is reasonable to assume that as specific failures accumulated, the principal source of threat became the possibility of failure on the entire *test* rather than on certain items. That total recall was significantly reduced under stress (compared to neutral conditions) also bears out this point. Thus, motivation may have determined selective recall of particular items during the early portion of the test, whereas recall of later items was largely a function of more fortuitous factors, such as impressiveness of content. Such effects would explain why there was some tendency for selective recall to relate to perseverance, but also why a demanding statistical technique such as Pearson's r would fail to yield a significant relationship.

2. Our findings have direct implications for the recent work of Alper, Eriksen, and Atkinson who report relationships between indices of hysteria, ego strength, need Achievement, and selective recall.

If our perseverance factor may be taken to include a stylistic component related to ego strength, then our findings of cognitive sensitivity to failure by high perseverance *S*s and insensitivity to failure by low-perseverance *S*s would seem to constitute a direct contradiction of Alper's (1957) results, where ego strength is related to superior recall of successes. The existence of a real contradiction assumes, of course, the comparability of the two indices of ego strength. Since our various tests designed to measure n Recognition, n Dominance, Counteractive Achievement, and Dejection Pessimism (all assumed by Alper to relate to ego strength) did not covary and loaded in separate orthogonal factors, Alper's index may indeed be multidimensional and only partially related to ego strength.

Our findings also stand in contradiction to those of Eriksen who, like Alper, reports a relationship between ego strength and superior recall of successes under stress. Since Eriksen's Rorschach indicator of ego strength failed to load on our perseverance factor and also shared little variance with the test battery, again the contradiction with our results would seem more apparent than real. It should be pointed out, however, that this indicator alone correlated nonsignificantly with both the CR-IR and PT measures (−.11 and .11 respectively). Likewise, both the Hysteria scale alone and the Psychaesthenia scale alone correlated nonsignificantly with these measures (.11 and −.18 respectively for the former and −.07 with both for the latter), although significant relationships would be expected from Eriksen's previous work.

The direction of association between the perseverance factor and our cognitive variables exactly parallels the findings of Atkinson (1953) and Atkinson and Raphelson (1956) with respect to need Achievement, thereby tending to support our hypothesis of overlap between these two dimensions.

3. Atkinson's data and our own thus suggest a clear relationship between a particular personality area and defense; namely, persons low in achievement concern and in achievement striving tend to repress the memory of failure experiences, whereas persons high in these factors tend to ruminate over these experiences and keep them cognitively focal. Three possible explanations may be advanced for this relationship. The first rests on the assumption that the low *S* is more fearful of failure than the high *S* and, hence, that greater anxiety is aroused for him as a consequence of the failure experience. This is the position taken by Atkinson (1953) with respect to low need achievers (high need achievers, he believes, are more "hopeful for success" and, although he is not explicit on this point, presumably they perseverate on the failure experience as part of an adaptive push for success). The second explanation is similar to the first in that it ties the repressive reaction of the low *S* to greater anxiety on his part, but it bases this anxiety on

unacceptable hostile impulses that have been cued off by the failure experience (hostility to the experimenter, to one's peers, or possibly even to oneself). This interpretation is more compatible with the classical psychoanalytic conceptualization of the repressive defense. At issue between these two positions is the question of whether negatively sanctioned *impulses* (presumably socialized during early infancy and leading to fear of *punishment*) or negatively valued states of *being* (which arise somewhat later as the individual develops a concept of himself and which lead to fear of *rejection* or of abandonment) are the precursors of the anxiety underlying repression.

The third explanation is suggested by our finding that the Neuroticism and Perseverance factors were orthogonal, and also that low- and high-perseverance *Ss* exhibited an equal pre- to poststress drop in rote memory for digits—both observations implying equal anxiety arousal in Lows and Highs. This interpretation therefore assumes *equal* fear of failure in Lows and Highs, but maintains that whereas the latter characteristically ruminate over, or cognitively converge upon, anxiety-inducing situations, the former characteristically repress them (as an extension of stylistic differences in persistence).[18] Thus, whereas the first two positions link repression to degree of anxiety arousal—i.e., to the *drive* property of anxiety—the third ties it to predisposed ways of handling anxiety—i.e., to the *cue* property of anxiety. Needless to say, there is as yet little empirical basis for choosing among these three alternatives.

SUMMARY

The present study had two objectives: (*a*) to isolate the particular personality dimensions that underlie indices in the conventionally delineated areas of ego strength, need Achievement, and hysteria; (*b*) to determine which of these dimensions relate to repressive and obsessive forms of defense. Thirty-eight questionnaire and performance indices were administered to 84 Harvard freshmen. The data were factor-analyzed, yielding five orthogonal dimensions tentatively labeled Neuroticism, Intellectual Flexibility, Extraversion, Other-orientation, and Perseverance for Achievement. Factorial scores for 42 *Ss* who underwent a Zeigarnik type failure-stress experience were correlated with two cognitive measures assumed to be indicators of defensive orientation: recall of task successes vs. failures, and perceptual recognition limens for test-related vs. neutral words. The correlation between Perseverance for Achievement and perceptual recognition was highly significant, Highs reacting obsessively and Lows repressively. The implication of the factor analysis for the concepts of *hysteria*, *need Achievement*, and *ego-strength* were discussed, and the correlations with recall and perceptual measures were examined, with special reference to studies by Alper, Eriksen, and Atkinson.

[18] See also Rosenzweig's (1943) discussion in this regard.

REFERENCES

ADORNO, T., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality.* New York: Harper, 1950.

ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: Correlation between experimental and personality data. *J. Pers.*, 1948, **17**, 104–137.

ALPER, THELMA G. Predicting the direction of selective recall: Its relation to ego strength and *n* achievement. *J. abnorm. soc. Psychol.*, 1957, **55**, 149–165.

ATKINSON, J. W. The achievement motive and recall of interrupted and completed tasks. *J. exp. Psychol.*, 1953, **46**, 381–390.

ATKINSON, J. W. *Motives in fantasy, action, and society.* New York: Van Nostrand, 1958.

ATKINSON, J. W., & RAPHELSON, A. C. Individual differences in motivation and behavior in particular situations. *J. Pers.*, 1956, **24**, 349–363.

CARON, A. J., & WALLACH, M. A. Recall of interrupted tasks under stress: A phenomenon of memory or of learning? *J. abnorm. soc. Psychol.*, 1957, **55**, 372–381.

CATTELL, R. B. Personality and motivation theory based on structural measurement. In J. L. McCary (Ed.), *Psychology of personality.* New York: Logos, 1956. Pp. 65–119.

EDWARDS, A. L. *Personal Preference Schedule.* New York: Psychological Corp., 1954.

ERIKSEN, C. W. Psychological defenses and "ego strength" in the recall of completed and incompleted tasks. *J. abnorm. soc. Psychol.*, 1954, **49**, 45–50.

EYSENCK, H. J. *The structure of human personality.* London: Methuen, 1953.

GUILFORD, J. P. *An inventory of factors STDCR.* Beverly Hills, Calif.: Sheridan Supply, 1940.

HATHAWAY, S. R., & MCKINLEY, J. C. *The Minnesota Multiphasic Personality Inventory.* New York: Psychological Corp., 1943.

HYMAN, R., & JENKIN, N. S. Involvement and set as determinants of behavioral stereotypy. *Psychol. Rep.*, 1956, **2**, 131–146.

KLEIN, G. S. The personal world through perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception —An approach to personality*. New York: Ronald, 1951.

KNUTSON, A. L. Personal security as related to station in life. *Psychol. Monogr.*, 1952, **66**, No. 4 (whole No. 336).

MCARTHUR, C. Personalities of public and private school boys. *Harv. educ. Rev.*, 1954, **24**, 256–262.

MCCLELLAND, D., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. *The achievement motive*. New York: Appleton-Century-Crofts, 1953.

MCCLELLAND, D., & LIBERMAN, A. M. The effect of need for achievement on recognition of need-related words. *J. Pers.*, 1949, **18**, 236–251.

MCREYNOLDS, P. Perception of Rorschach concepts as related to personality deviations. *J. abnorm. soc. Psychol.*, 1951, **46**, 131–141.

MURRAY, H. A. *Memorandum on pride*. Unpublished manuscript, Harvard Psychological Clinic, 1942. (Mimeo.)

ROSENZWEIG, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.

ROSENZWEIG, S., & SARASON, S. An experimental study of the triadic hypothesis: Reaction to frustration, ego-defense, and hypnotizability. I. Correlational approach. *Charac. & Pers.*, 1942, **11**, 1–19.

SNEDECOR, G. W. *Statistical methods*. (4th ed.) Ames, Ia.: State Coll. Press, 1946.

WHITE, R. W. Hypnotic test. In H. A. Murray (Ed.), *Explorations in personality*. New York: Oxford, 1938. Pp. 453–461.

Received July 11, 1958.

THE EFFECTS OF CHLORPROMAZINE ON GSR CONDITIONING[1, 2]

LONNIE E. MITCHELL

Veterans Administration Mental Hygiene Clinic, Washington, D. C.

AND MELVIN ZAX

University of Rochester

THE hypothesis tested in this study derives from two separate bodies of evidence. The first relates to the studies and clinical observations which have been done with one of the ataractic drugs, chlorpromazine or thorazine (the brand name of the Smith, Kline, and French laboratories). Those who have used it and observed its effects (Cowden, Zax, & Finney, 1955; Cowden, Zax, Hague, & Finney, 1956; Deloy, Deniker, & Harl, 1952; Lehmann & Hanrahan, 1954) report that it is particularly useful in quieting the disturbed, anxious patient.

Second, in a number of studies in a different area, Spence and his associates (Farber & Spence, 1953; Spence & Beecraft, 1954; Spence & Farber, 1953; Spence & Farber, 1954; Spence, Farber, & Taylor, 1954; Spence & Taylor, 1951) have demonstrated a positive relationship between the anxiety level of *S*, usually measured by the Taylor Manifest Anxiety Scale, and the ease with which *S* acquires a conditioned eyelid response. Other researchers (Bitterman & Holtzman, 1952; Welch & Kubis, 1947a; Welch & Kubis, 1947b) have worked with GSR as the conditioned response and reported similarly that anxious *S*s condition more readily.

From these lines of evidence, it seems appropriate to hypothesize that since thorazine appears to alter the anxiety level of a patient, and since conditionability seems related to anxiety level, taking this drug should result in a change in *S*'s capacity to acquire a conditioned response. More specifically, after having been administered thorazine for a period of time (in this case 30 days), *S* should condition with greater difficulty than was the case before. Similarly, drug levels should exert a differential effect on the conditioning measure.

METHOD

Sixty-nine psychiatric patients at St. Elizabeths Hospital in Washington, D. C., were used as *S*s. All met the following criteria: (*a*) they were newly admitted to the hospital; (*b*) they had not received thorazine medication for a period of at least four weeks prior to the initiation of the experimental procedures; (*c*) they were between 15 and 60 years of age; (*d*) they were not diagnosed as suffering from brain damage; (*e*) they were not mentally defective; (*f*) they had not undergone electroconvulsive shock, insulin coma therapy, or sodium amytal interviews for at least 30 days prior to the beginning of the experiment; and, finally, (*g*) they were able to cooperate in taking pencil-and-paper tests and in the conditioning situation. A total of 115 *S*s were referred for the study, but 46 were lost for the following reasons: (*a*) medical emergencies resulting in the drug's being withdrawn from some experimental *S*s or administered to some controls by the attending psychiatrists (involving 17 *S*s); (*b*) administrative errors resulting in *S*'s starting on a course of drug therapy before being assigned to any group (involving seven *S*s); (*c*) discharge or transfer from the hospital (involving 10 *S*s); (*d*) lack of cooperation (involving six *S*s); and (*e*) later diagnosis of organic involvement (involving six *S*s).

*S*s were divided between an experimental group of 36 and a control group of 33, equated for age, sex, length of hospitalization, and degree of anxiety as judged clinically by the referring psychiatrist.

After referral, *S*s were conditioned by *E*, following which all pertinent information about them was submitted to two collaborators who assigned them to either the experimental or the control group solely on the basis of the control variables. The collaborators then informed the psychiatrist in charge of the case whether the patient was to receive thorazine or not. They were also responsible for informing *E* at the end of approximately 30 days that it was time to recondition *S*. Not until all data were collected was *E* informed as to which *S*s belonged in which groups.

The dose level and method of thorazine administration were regulated by the psychiatrist in charge of the case. Usually, however, after a week or 10 days of gradual increase, most *S*s were receiving between 200 mg. and 600 mg. (orally) daily.

Conditioning took place in a semi-sound-proofed room where temperature was relatively uniform. The

[1] This paper is derived from a study done by the senior author in partial fulfillment of the requirements for the Ph.D. at American University. The data were collected at St. Elizabeths Hospital, and he is grateful to the administration and personnel, especially Albert D. Annis, for their cooperation.

[2] The authors wish to express their appreciation for the helpful comments made on the manuscript by Robert Ader and Eugene Sachs, both of the University of Rochester.

word "carry" was used as the CS for all *Ss* in the initial conditioning session. This word appeared nine times disbursed randomly among a list of 21 other commonly used five-letter words making a list of 30. Each word was presented in the aperture of a memory drum for an interval of 5½-seconds with an interval between successive stimuli of 1 sec. This word list was repeated as often as was necessary until conditioning was achieved. A different list of words and CS were used for the second conditioning session.

The UCS was the noxious sound of a bicycle horn of 1200 cycles per sec. in frequency and 50 decibels in intensity. It was delivered to *S*'s ear through a pair of high quality earphones attached to an audiometer (Maico Model H-I). The audiometer and earphones were used to establish equal loudness in both ears and to reduce the effect of any uncontrollable extraneous noises. The first appearance of the CS was always accompanied by UCS, but thereafter a 25% periodic reinforcement schedule was followed so that every fourth appearance of the CS was reinforced. On those trials when UCS and CS were paired, an electrical timing device automatically sounded the UCS 1⁹⁄₁₀-sec. after the CS appeared and interrupted it after ⁹⁄₁₀ of a second.

The CR was the GSR (resistance change) which accompanied the CS. This was measured by a Lafayette galvanometer using finger electrodes and a multichannel recorder (No. 603), which traced the GSR graphically and continuously throughout the session. Zinc oxide paste was applied as a conductor supplement. The criterion for conditioning was patterned after that of Welch and Kubis (1947a, 1947b). In this instance, it was two sets of three consecutive GSR reactions to the CS when unaccompanied by the UCS. A reaction was defined as an observable rise in the GSR accompanying the CS which was greater than that accompanying the word immediately preceding the CS. The number of reinforcements necessary before this criterion was achieved constituted the conditioning measure.

RESULTS

Table 1 presents data regarding the sex, age, diagnoses of *Ss*, and the number of reinforcements required to condition before and after the experimental period. In order to test the hypothesis that *Ss* undergoing thorazine therapy for 30 days would condition with greater difficulty than control *Ss* who failed to receive drug therapy, difference scores between number of trials to condition initially and number of trials required 30 days later were computed for all *Ss*. The scores of experimental *Ss* were then compared to those of controls by Fisher's *t* test for uncorrelated means. The experimental group required a mean of 6.78 more reinforcements to condition to the criterion; the standard deviation was 2.09. Controls required on the average .52 fewer trials to condition; the standard deviation in this case was 1.94. The *t* value of 15.21 is significant at far beyond the .001 level. Figure 1 graphically depicts the distribution of differences within each group and illustrates an effect which is so consistent that it results in distributions which are virtually nonoverlapping.

Since information was available on the amount of thorazine received by each patient, it was possible to examine the effects of variation in drug level on the change in conditioning

TABLE 1

SEX, AGE, AND DIAGNOSIS OF *Ss* AND GSR MEASURES BEFORE AND AFTER THE EXPERIMENTAL PERIOD

Experimental Group					Control Group				
Sex	Age	Diag.	GSR Pre	Reinf Post	Sex	Age	Diag.	GSR Pre	Reinf Post
M	43	S[a]	5	11	M	21	P	8	10
M	29	S	8	13	M	24	S	9	9
M	24	NP[b]	8	16	M	22	S	10	8
M	34	S	9	12	M	31	NP	10	9
M	27	S	9	14	M	38	S	10	9
M	32	S	9	18	M	42	NP	10	10
M	28	S	11	17	M	24	P	11	10
M	22	S	11	18	M	47	S	11	10
M	42	S	12	16	M	43	S	11	11
M	39	NP	13	17	M	46	S	11	11
M	43	S	13	19	M	16	NP	12	10
M	37	S	13	21	M	60	S	12	11
M	48	S	14	20	M	41	S	12	11
M	50	P[c]	14	27	M	31	S	12	16
M	19	NP	15	25	M	56	S	14	12
M	32	S	16	24	M	37	S	15	14
M	40	S	17	31	M	27	S	16	14
M	17	S	19	29	M	23	S	18	16
F	32	P	5	9	F	27	NP	4	5
F	29	S	6	12	F	48	P	4	7
F	29	S	6	12	F	33	S	6	6
F	23	S	6	12	F	26	NP	7	9
F	46	P	6	13	F	28	S	7	9
F	29	S	7	12	F	19	S	8	7
F	41	S	7	13	F	39	S	8	7
F	36	S	7	15	F	50	S	8	11
F	31	NP	8	13	F	52	S	10	4
F	34	S	9	14	F	36	S	10	9
F	60	NP	9	16	F	49	S	10	9
F	22	S	10	14	F	36	S	11	9
F	44	S	10	20	F	37	S	11	9
F	52	S	11	15	F	48	NP	11	10
F	42	P	11	19	F	33	P	16	14
F	37	S	12	19					
F	43	S	14	21					
F	48	NP	15	22					
Mean	35.67		10.42	17.19		35.88		10.39	9.88

[a] S represents a diagnosis of schizophrenic reaction.
[b] NP represents a diagnosis of psychoneurotic reaction.
[c] P represents a diagnosis of psychosis other than schizophrenia.

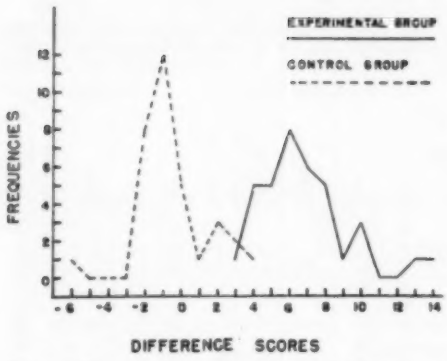

FIG. 1. DIFFERENCE SCORES BETWEEN REINFORCEMENTS REQUIRED FOR CONDITIONING BEFORE AND AFTER THE EXPERIMENTAL PERIOD

levels. The 36 experimental *S*s were divided into three groups according to whether they had received a high drug dosage (435 to 850 mg./day), a medium drug dosage (275 to 410 mg./day), or a low drug dosage (90 to 275 mg./day). The high drug group required an average of 6.00 more reinforcements, the medium group required 8.08 more reinforcements, and the low drug group required 6.25 more reinforcements before reconditioning. An analysis of variance yielded an F ratio of 4.05, significant at beyond the .05 level.

DISCUSSION

The findings of this study very clearly uphold the primary hypothesis that conditioning in humans is made more difficult by the administration of thorazine. This finding has implications both for the utility of the drug and for the value of the conditioning procedure as a situational measure of anxiety. Thorazine has been demonstrated to have marked effects on certain kinds of observed behaviors, but these have not always been clearly delineated and observed in a well controlled setting. The reported experimental procedure provided such a setting, and the predicted relationships were found to obtain.

In addition, one may reason that acquisition of a conditioned response is a function of the strength of drive and that, in the case of the noxious stimulus used here, the drive was anxiety. Thus, an impairment in *S*'s ability to acquire the conditioned response as the result of taking thorazine for a period of time suggests a reduction in drive affecting conditioning. This is in agreement with the clinical observation that thorazine produces a reduction in anxiety and tension.

The present findings also confirm at the human level those in studies by Ader and Clink (1957) and Miller, Murphy, and Mirsky (1957) with rats. The former found that animals treated with thorazine were inferior to controls in the acquisition of an avoidance response motivated by the fear of electric shock. Miller et al. (1957) worked with the same type of response and found that thorazine facilitated the extinction of such avoidance behavior.

The findings with regard to dosage levels suggest that moderate amounts of thorazine result in greater resistance to conditioning than high dosages. While these results are suggestive, they are probably useful only insofar as they offer leads which might be followed in a research program designed to investigate this variable with more directness and rigor. One of the obvious shortcomings of these data is the fact that the regrouping of *S*s which was done for the analysis failed to control for relevant factors like sex, age, and diagnosis.

SUMMARY

The effects of thorazine on the anxiety level of 36 psychiatric patients was investigated with GSR conditioning as the measure of anxiety. Thirty-three other psychiatric patients served as controls. Experimental *S*s were administered the drug for 30 days and were conditioned before and after this period. Controls were conditioned before and after a similar interval but received no drug. The groups were found significantly dissimilar in the differences between the number of reinforcements necessary to achieve conditioning before and after the 30-day period. The experimental group required more reinforcements to condition after receiving the drug than they had before having it.

The differences in number of reinforcements required to achieve conditioning before and after thorazine therapy were evaluated for those receiving high, medium, and low dosages. Those receiving only a moderate amount of thorazine seemed most difficult to recondition.

These data are viewed as merely suggestive and subject to a more rigorous study in which factors like age, sex, and diagnosis are better controlled in relation to dosage levels.

REFERENCES

ADER, R., & CLINK, D. W. Effects of chlorpromazine on the acquisition and extinction of an avoidance response in the rat. *J. Pharmacol. exp. Ther.*, 1957, **121**, 144–148.

BITTERMAN, M. E., & HOLTZMAN, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, **47**, 615–623.

COWDEN, R. C., ZAX, M., & FINNEY, R. C. A preliminary note on the use of chlorpromazine with neuropsychiatric disorders. *AMA Arch. Neurol. Psychiat.*, 1955, **73**, 700–701.

COWDEN, R. C., ZAX, M., HAGUE, J. R., & FINNEY, R. C. Chlorpromazine: Alone and as an adjunct to group psychotherapy in the treatment of psychiatric patients. *Amer. J. Psychiat.*, 1956, **112**, 898–902.

DELOY, J., DENIKER, P., & HARL, J. M. The treatment of excitement and agitation states by a method of medication derived from hibernotherapy. *Ann. med.-Psychol.*, 1952, **110**, 251–273.

FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.

LEHMANN, H. E., & HANRAHAN, G. E. Chlorpromazine: New inhibiting agent for psychomotor excitement and manic states. *AMA Arch. Neurol. Psychiat.*, 1954, **71**, 227–237.

MILLER, R. E., MURPHY, J. V., & MIRSKY, I. A. The effect of chlorpromazine on fear-motivated behavior in rats. *J. Pharmacol. exp. Ther.*, 1957, **120**, 379–387.

SPENCE, K. W., & BEECRAFT, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, **48**, 399–403.

SPENCE, K. W., & FARBER, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 116–119.

SPENCE, K. W., & FARBER, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, **47**, 127–134.

SPENCE, K. W., FARBER, I. E., & TAYLOR, ELAINE. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, **48**, 404–408.

SPENCE, K. W., & TAYLOR, JANET A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.

WELCH, L., & KUBIS, J. Conditioned PGR (psychogalvanic response) in states of pathological anxiety. *J. nerv. ment. Dis.*, 1947, **105**, 372–381. (a)

WELCH, L., & KUBIS, J. The effects of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83–91. (b)

Received July 14, 1958.

SELECTIVITY IN EXPOSURE TO INFORMATION

JUDSON MILLS[1]
U. S. Army Leadership Human Research Unit, HumRRO

ELLIOT ARONSON
Stanford University

AND HAL ROBINSON
San Francisco State College

A PERSON reading a newspaper does not devote the same amount of attention to each item he sees; he concentrates on some items and ignores others. We are apt to explain such selectivity in exposure to information by saying that people are more interested in some things than others. But to say this only raises the question: Why are they interested? What are the variables which determine "interest"?

One variable determining interest has been investigated in a recent study of post-decision exposure to information (Ehrlich, Guttman, Schönbach, & Mills, 1957). According to the hypothesis tested in this study, the degree of interest that information has for a person who has recently made an important decision depends on whether or not it supports the choice he has made. Persons tend to seek out material that they expect to favor their choice, and avoid information that seems likely to be discrepant with the action they have taken. This hypothesis was derived from Festinger's (1958) theory of cognitive dissonance. The study found that recent purchasers of new automobiles read advertisements of their own car more often than advertisements of other new cars. Naturally, this study dealt only with positive information, that is, information presenting good features of an alternative; advertisements never mention faults of the product advertised.

However, the hypothesis is not limited to information that is positive; it applies equally well to negative information about the alternatives. Just as a person should tend to seek out information which he expects to contain good features of the alternative he has chosen and should avoid positive information about the rejected alternative, he should also tend to avoid information which he expects to give bad features of the chosen alternative and should seek out information that contains negative features of the alternative he rejected. Thus, for example, new car owners should avoid reading material giving faults of their own car and should prefer material that contains faults of other cars. One purpose of the present study was to test the implications of the hypothesis for exposure to both positive and negative information.

A second purpose was to determine the effect of the importance of the decision upon subsequent selectivity in exposure to information. According to the theory of cognitive dissonance, the more important the decision is, the greater is the post-decision selectivity.

To summarize, these hypotheses were tested in the present study:

1. Following a decision, persons tend to seek out information that favors the chosen alternative and to avoid information that favors the rejected alternative. That is, they prefer positive information about the chosen alternative to positive information about the rejected alternative, and prefer negative information about the rejected alternative to negative information about the one chosen.

2. The more important the decision, the greater is the subsequent selectivity in exposure to information.

METHOD

The method used to test the hypotheses had two essential features. First, college students were given their choice of the kind of examination they were to take. Second, after they had made a decision, the students were presented with a list of articles about the two alternative types of exams and were asked to indicate which articles they preferred to read.

The experiment was run under four different conditions; High importance, Positive information; Low importance, Positive information; High importance,

[1] Now at Syracuse University.

Negative information; and Low importance, Negative information. There were three sections of an introductory psychology course in each condition.[2] All 12 of the classes were seen in a single morning, with four persons playing the role of the experimenter (*E*).[3] The conditions were balanced with regard to hour of day and person acting as *E*.

The *E* introduced himself as a representative of the Educational Practices Committee of the Psychology Department. He said that the instructor had given him permission to meet with the class to tell them what the committee had been doing. The *E* explained that the committee was going to try out something new in some of the psychology classes. It had been suggested that students would do better when they were given the choice of the kind of examination that they would take. They were going to test this idea and actually give students their choice of the kind of exam they would take. The *S*s were given a choice of either an essay examination or an objective examination. It was stressed that each person would receive the examination of his choice and that once he made a decision he would not be able to change it.

In the High importance conditions, the examination was to be the final for the course, and the *S*s were told that, "Because you are being given your choice, we are going to count this quite a bit. Specifically, it will count 70% of your total grade for this course." *S*s in the Low importance conditions were told that, "Because this is a new thing, we aren't going to count this examination very much. Actually, it won't be more than a quiz and it will count only 5% of your total grade."

After the *S*s' choices were collected, the *E* explained that the instructor had felt the students should become acquainted with the work psychologists have done concerning examinations. Therefore, the instructor was going to have them read one scientific article dealing with examinations. The *S*s were told that copies of six different articles were available but since there was a limited supply of each article it might not be possible to give them their first choice, or perhaps not even their second choice. The *S*s were given a list of six fictitious articles and were asked to rank them in terms of how much they preferred to read them, placing a figure 1 next to the article they would like most to read, a 2 next to the article they would like next most to read, etc. Three of the articles on the list were about essay examinations, and the other three were about objective examinations. Each article was preceded by a one-sentence description. In the Positive information conditions, the articles were described as presenting arguments in favor of the kind of examination the article was about. In the Negative information conditions, the key words in the description were changed so that the same articles with identical authors, titles, journals, and dates were described as presenting arguments against the kind of examination the article was about.

After the *S*s had ranked the articles, and the sheets had been collected, the *E* spent the remainder of the hour explaining the hypothesis and how the experiment was designed to test it. The deceptions were revealed and the reasons for them discussed. The students were cautioned not to say anything about the experiment until the day was over.

RESULTS AND DISCUSSION

The degree to which a student preferred to read articles about one kind of examination more than articles about the other kind is reflected in the rankings of the two kinds of articles, the type of article preferred receiving the smallest ranks. Thus, the sum of the ranks of the three articles concerning the chosen examination provides a measure of selectivity in exposure to information. This measure has a possible range from 6 to 15. Since the sum of the ranks for all six articles is 21, the average for this measure should be 10.5 when the articles about the two kinds of examinations are equally preferred. (Of course, for each individual this measure can only be a whole number.)

From the hypotheses we should expect that the articles should not be equally preferred. The average of the sum of the ranks of the three articles concerning the chosen examination should be smaller than 10.5 in the Positive information conditions, and smaller in the High importance Positive information condition than in the Low importance Positive information condition. In the Negative information conditions, the mean of the sum of the ranks of the three articles concerning the chosen examination should be greater than 10.5

TABLE 1

MEANS FOR THE SUM OF THE RANKS OF THE THREE ARTICLES ABOUT THE CHOSEN EXAMINATION

Choice	Importance	Information	
		Positive	Negative
Objective	High	8.74 (58)	10.86 (59)
	Low	8.82 (66)	11.41 (61)
Essay	High	7.58 (33)	9.83 (24)
	Low	7.20 (25)	9.60 (30)
Both	High	8.32 (91)	10.56 (83)
	Low	8.37 (91)	10.83 (91)
Both	Both	8.35 (182)	10.71 (174)

Note:—*N*s given in parentheses.

[2] The authors are grateful to the psychology faculty of San Francisco State College for making the classes available for the experiment.

[3] Stewart Oscamp deserves thanks for acting as an experimenter.

and should be greater in the High importance Negative information condition than in the Low importance Negative information condition.

Table 1 presents the means of the sum of the ranks of the three articles about the chosen examination for those who chose essay examinations and for those who chose objective examinations in the four different experimental conditions.[4] In each condition, approximately two-thirds of the *Ss* chose the objective examination. An examination of Table 1 reveals that the mean for those in the Positive information conditions is, as expected, considerably less than 10.5, specifically, 8.35. The difference is significant beyond the .001 level. Contrary to the prediction, the mean for the High importance Positive information condition is not smaller than the mean for the Low importance Positive information condition. Also contrary to prediction, the mean for those in the Negative information conditions is very nearly equal to 10.5, which means that the articles about the two kinds of examinations were equally preferred in these conditions. And the mean for the High importance Negative information condition is not greater than that for the Low importance Negative information condition. It should be noted that the results are consistent no matter which kind of examination was chosen.

On the basis of these results, what can we say about the hypotheses? The data certainly do not support the prediction about the importance of the decision. Although it would seem that the degree to which an examination determines the course grade should influence the importance of a decision concerning that examination, it is possible that the procedure did not create differences in the importance of the decision. It is conceivable that any decision concerning examinations is highly important to college students. However, there is evidence that the manipulation of importance was effective. After they had made their choice, the *Ss* were asked to rate how difficult it was to make the decision on a six-point scale from "not at all difficult" to "extremely difficult." *Ss* in the High importance conditions rated the decision as significantly more difficult than did those in the Low importance conditions ($p < .02$). We must conclude that the importance of the decision did not influence selectivity in exposure to information as predicted by dissonance theory.

On first examination, it seems that the data also fail to support the hypothesis that persons seek out information supporting their choice, and avoid information discrepant with it. There was no selectivity in the Negative information conditions. It is true that selectivity did occur with respect to positive information. One might attempt to explain the results in the Positive information conditions by proposing that people want to find out more about the alternative they have chosen. This hypothesis seems plausible and would also explain the finding reported previously (Ehrlich et al., 1957) that new car owners read more advertisements about their own cars. But this explanation cannot account for the results in the Negative information conditions. According to this hypothesis, negative information about the chosen alternative should be preferred more than negative information about the rejected alternative. But this was not the case; there was no selectivity in the Negative information conditions.

However, if we use both of these hypotheses, (*a*) that persons tend to seek out information that supports their choice and to avoid discrepant information, and (*b*) that persons seek more information about the thing they have chosen, we can explain the data for both positive and negative information. In the Positive information conditions, both these effects should have operated in the same direction, and it was in these conditions that the students showed a strong tendency to prefer articles about the chosen examination. In the Negative information conditions, these two effects should have worked in opposite directions, and could have cancelled each other so that selectivity did not occur.

SUMMARY

College students were given their choice of the kind of examination they were to take. The importance of the decision was varied by telling some groups that the examination would count very much toward their course

[4] Tabulation of the data was facilitated by funds from a National Science Foundation grant administered by Leon Festinger.

grade and other groups that it would count very little. After they had made a decision, they were presented with a list of fictitious articles about the two alternative types of examinations and were asked to indicate which articles they preferred to read. For some groups, each of the articles appeared to contain positive information about one or the other of the two kinds of examinations; for other groups, each article appeared to contain negative information. When the information was positive, articles about the chosen examination were preferred. When the information was negative, articles about the two kinds of examination were equally preferred. The importance of the decision did not influence selectivity. The results were explained by the combined operation of two tendencies; (*a*) that persons tend to seek out information that supports their choice and to avoid discrepant information, and (*b*) that persons seek more information about the alternative they have chosen.

REFERENCES

EHRLICH, D., GUTTMAN, I., SCHÖNBACH, P., & MILLS, J. Post decision exposure to relevant information. *J. abnorm. soc. Psychol.*, 1957, **54**, 98–102.

FESTINGER, L. *A theory of cognitive dissonance.* Evanston, Ill.: Row, Peterson, 1958.

Received July 14, 1958.

THE EFFECT OF VERBAL REINFORCEMENT ON THE RECALL OF EARLY MEMORIES[1]

HERBERT QUAY

Vanderbilt University

Most present-day theories of human personality and of the historical and dynamic causes of human behavior were originally formulated by individuals practicing psychotherapy. In fact, it was the memories and free associations of neurotic patients that led Freud to formulate his theory of psychosexual development. Similarly, the memories and free associations of patients talking to other such innovators have led to very different conceptions of the origins of human personality. Even though these personality theorists presumably began with the same raw verbal materials, however, it is apparent that they derived quite different formulations of human motives and conflicts. Moreover, the various schools have defended their positions mainly on the basis of further psychotherapeutic contacts with patients, with no resolution of the theoretical differences. One may suspect, therefore, that observations which are highly stimulating in the formation of hypotheses may not be very useful in their verification or testing (Hunt, 1952).

Logical analysis makes it evident that if psychoanalytic therapy, or any other form of psychotherapy, is to serve the hypothesis-testing as well as the hypothesis-finding role, it is absolutely necessary to assume that the verbal productions of patients are free of influence from the therapist. Suspicions that the psychotherapist, however unintentionally, does influence subtly the productions of his patients have already been voiced. Bach (1952) observed that the human figure drawings produced by his patients during psychotherapy were different in their content from those produced by patients of his psychoanalyst officemate. Furthermore, the differences appeared in the way the content reflected the theoretical biases of each therapist. Heine (1953) studied patients who had successfully completed psychotherapy with therapists of different theoretical orientations. He found that the reasons given by patients for their improvement corresponded to the theoretical convictions of their therapists about what is therapeutic.

The recent literature on the operant conditioning of verbal behavior may provide a paradigm to account for these subtle influences. Stemming from the work of Greenspoon (1951, 1955), Verplanck (1955a, 1955b), and Wilson and Verplanck (1956), a rather wide variety of verbal behavior has shown itself subject to manipulation by another person through the use of verbal reinforcing stimuli. From the increase in the probability of saying words of certain classes (Greenspoon: 1951, 1955; Sidowski, 1954; Wilson & Verplanck, 1956) through the increase in complex types of discourse (Verplanck, 1955a), a wide variety of verbal behavior has been manipulated by the use of reinforcing stimuli with the added finding that subjects (*S*s) are generally unaware that their behavior has been manipulated. No study to date, however, has demonstrated that highly personal, emotionally charged verbalizations are subject to manipulation when the participation of the experimenter (*E*) is minimal. Since these are the types of verbal productions important in psychotherapy, it is necessary, for the argument presented here, to show that such verbal responses are subject to the effect reported for more impersonal materials.

The hypothesis tested in this study is that a verbal reinforcing stimulus increases the frequency of certain types of early childhood memories over their frequency in a "free-operant" period before any reinforcement is applied.

METHOD

The method of this study was to have a number of college students report memories of their early childhood. After a baseline was established in a "free-operant" period, verbal reinforcement at a minimal

[1] An abstract of a Ph.D. thesis done at the University of Illinois under the direction of J. McV. Hunt. The author wishes to express his appreciation to J. McV. Hunt, Lee J. Cronbach, G. R. Grice, L. A. Hellmer, D. R. Peterson, and M. M. Ohlsen, members of his doctoral committee.

level, the interviewer saying "uh-huh," was introduced to reinforce selected memory topics.

Subjects

Ss were chosen from a large class of undergraduates in introductory psychology at the University of Illinois, where each student is required to serve for five hours as *S* of psychological experimentation. *Ss* for this study volunteered for an experiment presented as involving the recall of childhood events.

The experimental groups were composed of 16 *Ss* in a group for which memories concerned with their families were reinforced ("family" group—nine females and seven males) and 18 *Ss* in a group for which memories concerned with people and events outside the family were reinforced ("nonfamily" group—ten females and eight males).

The "family" and "nonfamily" classes of memories were chosen because they could be identified reliably and could provide adequate initial operant levels to insure an early appearance of the operant response in the reinforcement period. Instructions for scoring, with examples, may be found in (Quay, 1958)

The reinforcing stimulus chosen was a flat, noncommittal "uh-huh," spoken in a low conversational tone. Such minimal reinforcement was chosen because it is in keeping with a general psychotherapeutic type of response and because if a response this minimal could influence verbal behavior, then almost any remark of a psychotherapist would be reinforcing.

All *Ss* reported for the experiment to *E*'s office at the Psychological Clinic. *Ss* were instructed to recall events from their early childhood as they came to mind. *Ss* were seated in a comfortable chair facing away from *E*. All responses were tape recorded by means of a microphone given to each *S* at the time he received his instructions.

The experimental period was divided into two half-hour sessions, generally one week apart. The first 10 min. of the first half-hour were used to establish the operant level (relative frequency) of the category subsequently to be reinforced. This procedure served to make each *S* his own control. At the end of the first half-hour, *S* was interrupted and asked to return one week later for another half-hour. At the end of the second and final half-hour, *Ss* were asked to cooperate further by filling out a questionnaire to be sent to them at a later date.

The questionnaire was designed to elicit information about the degree of *S*'s awareness of the contigency between his memories and the reinforcing stimulus. While it may have been more informative to question each *S* verbally at the end of the final half-hour, this was not done for fear that the real purpose of the study might become known to *Ss* who had not yet participated. After all *Ss* had completed the experiment, the questionnaire was mailed to each with a covering letter.

RESULTS

Reliability of Memory Categories

The reliability of the classification of memories by *E* when tested by a second judge showed a percentage of agreement that ranged between 90% and 100%, with a mean of 96%. The reliability was based upon a tape-recorded sample of 20 memories from each of 11 *Ss* taken as every third *S* from the combined groups. Reinforcements were eliminated from this tape to avoid presenting clues to the second judge.

TABLE 1
RELATIVE FREQUENCY OF MEMORIES IN THE REINFORCED CATEGORY FOR THE OPERANT PERIOD AND FOR THE REINFORCEMENT PERIOD FOR THE TWO EXPERIMENTAL GROUPS

"Family" Group				"Nonfamily" Group			
S	Operant	Reinforcement	Change	*S*	Operant	Reinforcement	Change
1	.30	.43	+.13	1	.30	.47	+.17
2	.39	.41	+.02	2	.57	.32	+.25
3	.37	.29	−.08	3	.81	.89	+.08
4	.40	.48	+.08	4	.78	.80	+.02
5	.25	.12	−.13	5	.60	.70	+.10
6	.55	.58	+.03	6	.53	.84	+.31
7	.50	.56	+.06	7	.90	.73	−.17
8	.06	.26	+.20	8	.60	.63	+.03
9	.60	.70	+.10	9	.69	.70	+.01
10	.50	.46	+.04	10	.58	.64	+.06
11	.50	.62	+.12	11	.63	.70	+.07
12	.33	.47	+.14	12	.35	.58	+.23
13	.53	.62	+.08	13	.50	.58	+.08
14	.53	.70	+.17	14	.55	.80	+.25
15	.38	.43	+.05	15	.45	.61	+.16
16	.33	.48	+.15	16	.48	.54	+.06
				17	.18	.33	+.15
				18	.58	.57	−.01

Comparison of Operant and Reinforcement Periods

To test the hypothesis that the reinforcing stimulus would increase the proportion of memories in the reinforced category, "family" or "nonfamily," the relative frequency in the operant period was compared to the relative frequency in the reinforcement period. This latter period was defined as that part of the total experimental hour which followed the delivery of the second reinforcement. Table 1 presents the relative frequency of the reinforced category (reinforced memories divided by all memories) for the operant period and for the reinforcement period of each *S*.

For the "family" group, 13 of the 16 *Ss* gave a higher proportion of "family" memories after reinforcement. The exact probability of

this result for a one-tailed sign test is .01064 (Siegel, 1956). The use of a one-tailed test is appropriate, since the direction of the change is predicted. For the "nonfamily" group, 15 of 18 *S*s gave a higher proportion of "nonfamily" memories after reinforcement. This result has an exact probability of .00377 for a one-tailed sign test.

For the groups combined, a technique suggested by Cronbach (1949, p. 411) was employed. For both groups, the relative frequency of "family" memories in the reinforcement period was plotted as a function of the relative frequency of "family" memories in the operant period. A median was obtained visually for *each* operant level class interval along the baseline and cases in each group falling above and below this median were noted. With these above and below the median counts combined into a single fourfold table (as if the count had been made in terms of a grand median for *all* cases regardless of operant level), chi square was computed. The obtained value was 12.44, significant at the .001 level for one degree of freedom. Inspection of the scatter plot and the fourfold table indicated that it was cases in the "family" group which more often were above the column medians.

Here it should be noted that while both groups are experimental groups, the mutually exclusive, all-inclusive nature of the memory categories makes them at the same time control groups for each other. For example, if one should argue that the relative frequency of "family" memories would naturally rise as a result of *S* becoming more at ease with *E* as the interview proceeded, then one must explain the drop in relative frequency of "family" memories when "nonfamily" memories were reinforced.

Magnitude of Effect

Inspection of Table 1 indicates further that there were differences in amount of change for those *S*s who showed an increase in relative frequency of the reinforced category. Increases ranged from .01 to .31. Since initial operant levels were different for most all of the *S*s, no exact comparisons of these changes are possible. However, no significant Mann-Whitney *U* tests (Siegel, 1956) suggest that these differences are unrelated to either "family" or "nonfamily" group or to initial operant level.

Awareness of Subjects

Of the 30 *S*s who returned the questionnaire regarding awareness, only one individual indicated that he was aware of the contingency between his responses and the reinforcing stimulus. This *S* ranked second in the magnitude of increase in relative frequency under reinforcement. Fifteen of the *S*s indicated that they recognized that *E* had said "uh-huh," yet all 15 denied having been influenced by *E*. Three *S*s indicated that *E* seemed more interested in some topics than others; but none named the topic, and all three denied influence. Eleven of the *S*s apparently had no awareness whatsoever of either *E*'s saying "uh-huh" or of any relationship between this and what they recalled.

DISCUSSION

This study has demonstrated that highly personal and emotionally charged memories can be manipulated by another person in an interview with very minimal verbal participation selectively placed after certain classes of these memories.

While this process of subtle interpersonal influence has not been demonstrated within psychotherapy itself, several considerations lead one to predict that the effect would appear with even greater magnitude over a series of psychotherapeutic interviews. For example, a psychotherapist would have much more time in which to respond selectively, his verbal participation in psychotherapy would be considerably greater than the very minimal reinforcing verbalization used in this study, and it seems very likely that he would have more prestige (Verplanck, 1955b) with his patients than the graduate-student investigator in the role of *E* had with the *S*s in this experiment.

Certain implications for the evaluation of psychotherapeutic success or failure are also suggested by the results of this study. One major point of emphasis in the evaluation of patient improvement has been the changes in the patient's verbal behavior over the course of treatment (Mowrer, Hunt, & Kogan, 1953; Raimy, 1948; Seeman, 1949). While such changes are assumed to be attributable to a restructuring of the personality as a result of insights gained in the psychotherapy, there

certainly exists the possibility that a psychotherapist may selectively reinforce these types of statements, increasing their frequency in the later stages of treatment. It is conceivable, however, that such changes of verbalization may constitute genuine therapeutic restructuring, but the validity of such a contention must be tested before it can be accepted.

The question of awareness is an intriguing one in relation to both the patient and the therapist. Many *S*s are able to learn in the verbal conditioning situation without verbalizable awareness of either the presence of the reinforcing stimulus, the response considered "correct," or the contingency between the two. That the therapist may selectively respond outside of his own awareness has already been suggested by Bach (1952). If research within psychotherapy itself demonstrates selective therapist reinforcement of patient verbalizations which make these verbalizations conform more and more to the therapist's theoretical convictions, one might well expect that this process was taking place outside the awareness of both patient and therapist.

SUMMARY

Thirty-four college students were asked to recall events from their early childhood. After a baseline was established in an initial operant period, 16 of the *S*s were reinforced by *E* saying "uh-huh" for memories concerned with members of the *S*'s family. Eighteen of the *S*s were similarly reinforced for memories not concerned with the family. In both groups the reinforcing stimulus served to increase the proportion of memories in the reinforced category when the reinforcement period was compared to the operant period. The results, indicating that personally relevant and emotionally changed materials are subject to manipulation by the use of a minimal reinforcing verbalization, were related to the testing of causal and dynamic hypotheses about personality within the psychotherapeutic interview.

REFERENCES

BACH, G. R. Some diadic functions of childhood memories. *J. Psychol.*, 1952, **33**, 87–98.

CRONBACH, L. J. Statistical methods applied to Rorschach scores: A review. *Psychol. Bull.*, 1949, **46**, 393–429.

GREENSPOON, J. The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response classes. Unpublished doctoral dissertation, Univer. of Indiana, 1951.

GREENSPOON, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409–416.

HEINE, R. W. A comparison of patient's reports on psychotherapeutic experience with psychoanalytic, non-directive, and Adlerian therapists. *Amer. J. Psychother.*, 1953, **7**, 16–23.

HUNT, J. McV. Psychological services in the tactics of psychological science. *Amer. Psychologist*, 1952, **11**, 608–622.

MOWRER, O. H., HUNT, J. McV., & KOGAN, L. S. Further studies utilizing the discomfort-relief quotient. In O. H. Mowrer (Ed.), *Psychotherapy theory and research*. New York: Ronald, 1953. Pp. 257–295.

QUAY, H. C. The effect of verbal reinforcement on the recall of early memories. Unpublished doctoral dissertation, Univer. of Illinois, 1958.

RAIMY, V. Self-reference in counseling interviews. *J. consult. Psychol.*, 1948, **12**, 153–163.

SEEMAN, J. A study of the process of nondirective therapy. *J. consult. Psychol.*, 1949, **13**, 157–168.

SIDOWSKI, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. exp. Psychol.*, 1954, **49**, 355–360.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

VERPLANCK, W. S. The control of the content of conversation; reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **51**, 668–676. (a)

VERPLANCK, W. S. The operant conditioning of human motor behavior. *Psychol. Bull.*, 1955, **53**, 70–82. (b)

WILSON, W. C., & VERPLANCK, W. S. Some observations on the reinforcement of verbal operants. *Amer. J. Psychol.*, 1956, **64**, 448 451.

Received July 14, 1958.

THE EFFECTS OF FATHER-ABSENCE ON NORWEGIAN BOYS AND GIRLS[1]

DAVID B. LYNN AND WILLIAM L. SAWREY
School of Medicine, University of Colorado

THE effects of the father on the development of the child may be approached by comparing children whose fathers have been absent from the home for long periods of time with those whose fathers have been present. Previous studies have been concerned with temporary absence of the father during wartime (Bach, 1946; Sears, R., Pintler, M. H. & Sears, P., 1946). In Norwegian sailor families, however, the father is often away from home for periods of two years or more with only infrequent and irregular periods at home. This paper is concerned with the separate and differential effects of father-absence on boys and girls in such Norwegian sailor families.

In a wartime study of six- through ten-year-old children Bach (1946) found that father-fantasies of father-absent boys were more similar to those of girls than were the fantasies of father-present boys. There was no comparable difference between the two groups of girls. Sears et al. (1946), in a doll play study of the effects of father-absence in wartime on preschool age children, found that father-absent boys were less frequently aggressive than father-present boys. No such finding was reported for the girls.

Tiller (1958), in an earlier report employing the same subjects on which this paper is based, compared Norwegian sailor families with appropriate control families in which the father was present (see below). According to his findings, mothers whose husbands were absent tended to have the following characteristics as compared with the control mothers: (*a*) a smaller proportion led active social lives; (*b*) a smaller proportion worked outside the home; (*c*) a larger proportion tended to be overprotective; and (*d*) a larger proportion tended to stress obedience and politeness as contrasted to happiness and self-realization for their children. Tiller reported that the father-absent children tended to have the following characteristics as compared with the control children: (*a*) A larger proportion showed a high degree of dependence; (*b*) a larger proportion showed pseudo-maturity; and (*c*) a larger proportion idealized the father.

Tiller considered his findings as resulting primarily from indirect effects of father-absence, mediated by the fact that father-absence for the child also means husband-absence for the mother. In his view, the mother's reaction to husband-absence is reflected in her treatment of the child, and this treatment in turn affects the child.

The present paper, as contrasted with Tiller's, is concerned with the separate and differential effects of father-absence on boys and girls rather than the general effects of father-absence on boys and girls combined. It was predicted that the direct effects of father-absence per se would affect boys and girls differentially. The direct effects of father-absence were hypothesized to be more damaging to the boy's development than to the girl's. The father-absent girl still has the same-sex parent with her as a model for identification, whereas the father-absent boy is virtually without the same-sex parent to serve as a model.

The following predictions were made as to the direct effects of father-absence: 1. More father-absent boys would show immaturity than father-present boys; 2. Being insecure in their identification with the father, the father-absent boys would show stronger strivings toward father-identification than father-pres-

[1] The data for the research project, of which this study is a part, were collected by Per Olav Tiller and the senior author of this paper at the Institute for Social Research, Oslo, Norway in 1956. Financial support was provided by a grant from a Norwegian shipowner, Anders Jahre. Tiller (1958) analyzed part of the data and published his analysis elsewhere. The data used in the present study were analyzed by the two present authors. We wish to thank Herbert Gerjuoy for his consultation. The senior author wishes to thank the Indiana University School of Medicine for granting a year's leave of absence to participate in this project.

ent boys; 3. The father-absent boys would react to their insecure masculine-identification with compensatory masculinity; and 4. The father-absent boys would demonstrate poorer peer adjustment than father-present boys and than father-absent girls. 5. It was also predicted that father-absent girls, perhaps threatened by the absence of one parent, would become more dependent on the mother than would father-present girls.

METHOD

Subjects

The *Ss* consisted of 80 mother-child pairs from several neighboring small towns and their environs in a typical sailor district in Norway. There were 40 mother-child pairs in the father-absent (experimental) group (21 boys and their mothers, 19 girls and their mothers). There were also 40 mother-child pairs in the father-present (control) group (20 boys and their mothers, 20 girls and their mothers). The children, all drawn from the second grade of the public schools, ranged from eight years of age through nine years six months. The two groups were comparable in the teachers' ratings of intelligence. The fathers, in the father-absent group, were sailors and whalers who were away from home at least nine months out of the year, and sometimes as long as two years or more, with only infrequent and irregular periods at home. The experimental families were restricted to those of officers, because only officers were consistently absent from their families for such long periods. The control group was selected so that the father's occupation would be as comparable as possible to that of whaler or sailor officer. It included only families where the father, like officers, supervised others, e.g., shop owners, high civil servants, and office managers.

From an analysis of data on the background characteristics of the present *Ss*, viz., urban–rural, age of mother, housing conditions, number of siblings, sibling order, and teachers' ratings of intelligence, Tiller (1958, p. 42) concludes "... that the differences in background data between our two samples do not explain the psychological differences between the groups." Nevertheless, the reader should bear in mind that unknown cultural factors specific to the samples may be operating instead of or in addition to the variable of father-absence. Women who marry sailors may, for example, be such a select group that the variable "sailor wife" may be the crucial one. Moreover, one cannot say whether such findings would hold with a comparable American population. It should be noted, however, that the hypotheses tested were derived from a theoretical formulation concerning the effects of father-absence generally, not of sailor families nor Norwegian families in particular.

Procedures

The mothers were administered a semistructured interview and the children the Structured Doll Play test. The children were also asked to draw a picture of a family.

The interview consisted of about 60 questions designed to elicit information about: (*a*) The mother-child relationship; (*b*) the father-child relationship; (*c*) the child's maturity level; and (*d*) his peer adjustment. There were questions such as, "Does the child often come to you with his troubles?" "Has he talked about wishing to be like his father?" "Does he often get into fights?" After the mother responded to a given question she was asked to, "Tell me about it," or "What are your feelings about that?"

The Structured Doll Play test (SDP) is a modification of a projective test developed by Lynn (1955),[2] which presents the *S* with doll-house dolls in a series of typical family and age-mate situations. Each situation is structured by telling the child what is going on, and the child then resolves the situation through doll play. For example, in the Crib or Bed Choice situation *S* is presented with doll-house crib, bed, and child doll (the sex of the child doll being the same as *S*). The *S* is told, "The little boy (girl) has two places he can sleep. He can sleep in either the crib or the bed. You put him to sleep." *S* is then asked, "Who comes in, Mommy or Daddy?" and, "Then what happens?" In the Toilet situation, a bathroom setting is placed before *S*. In this situation *S* is asked, "Which one takes him to the toilet, Mommy or Daddy?" He is then asked, "Which one does Mommy (Daddy, depending on the choice) take him to, the toilet or potty chair?" There are also eating, sleeping, fighting situations, etc.

The modified SDP used in this study consisted of 18 situations. One especially designed for this study involved telling *S* that the father doll went away for a long time, and then asking, "What happens while Daddy is gone?" After *S* had responded to the above instructions, the father doll was introduced and *S* was asked to tell what happens when the father returns.

Scoring Methods

All of the scoring was done by a technician who had no knowledge of the nature of the study or its hypotheses. The various rating scales are given below, with notes on the scoring criterion.

Crib or bed choice (SDP). The Crib or Bed Choice situation was coded as to whether *S* chose to have the ego doll sleep in the crib or the bed. The ego doll was the child doll used by a given *S* throughout the SDP.

Parent doll choice (SDP). This is a measure of the frequency with which the mother or father doll was chosen in the SDP. There were 10 situations in which *S* could choose either the mother or father doll. Low

[2] The original idea for this test grew out of a discussion with Bertram D. Cohen, Lafayette Clinic, Detroit, Michigan. B. D. Cohen also collaborated in the early stages of its development.

TABLE 1

FREQUENCY DATA FOR SIX VARIABLES CONCERNING FATHER-ABSENT (FA) AND FATHER-PRESENT (FP) NORWEGIAN BOYS AND GIRLS

Group	A Crib or Bed Choice (SDP)		B Father Choice (SDP)		C Compensatory Masculinity (Interview)	
	Crib	Bed	0-4	5-10	Low	High
FA boys	19	2	6	14	7	14
FP boys	11	9	13	7	17	3
FA girls	16	3	13	6		
FP girls	14	6	16	3		

Group	D Peer Adjustment (Interview)		E Separation (SDP)		F Dependence (SDP)	
	Poor	Adequate	Mother	Father	Low	High
FA boys	11	10	9	12	13	8
FP boys	3	17	11	9	16	4
FA girls	4	15	18	1	10	9
FP girls	5	15	10	10	18	2

Note:—The total frequencies in the individual tables are not always 80 because of difficulties beyond *E*'s control, e.g., lack of cooperation or incomplete responses.

father choice was considered to be from zero to four choices, and high father choice from 5 to 10 choices.

Compensatory masculinity (interview). Compensatory masculinity was coded as to its presence or absence in the interview. Statements in the interview suggesting that the boy behaves in an exaggeratedly masculine way were considered to be compensatory masculine if there were also suggestions in the same interview that the boy, at other times, behaves in a way more characteristic of girls than is usual for boys.

Peer adjustment (interview). The interview was rated on a four-point scale as to the peer adjustment of the child. The basis of judgment was the mother's statements concerning the number of friends the child had, his preference for playmates in his own or a different age group, his aggressiveness or the inhibition of aggression, or the clear indication of other problems with peers. In analysis, scores were dichotomized, with ratings of 1 and 2 assigned to the adequate adjustment category and ratings of 3 and 4 to the poor adjustment category.

Parent doll choice in the Separation situation (SDP). The Separation situation was coded as to whether the mother or father doll was chosen to return to the ego doll first following separation.

Dependence (SDP). The SDP was rated on a five-point dependence scale, based on responses in the three SDP situations concerning separation from one or both parent dolls. A response was scored high on dependence if the ego doll was shown to cry upon separation or to run after the parent doll(s). It was scored low if the ego doll was shown to face separation with constructive adjustment, such as playing alone, going over to a friend's house, or preparing himself a meal. The first three points were combined to form the low dependence category, and points four and five to form the high dependence category.

Draw-a-family. The *S*s were asked to draw a picture of a family. The family drawing was scored for three variables, viz., the parent figure drawn first, the one drawn largest, and the one drawn in most detail.

RESULTS

Chi square was used in all statistical analyses. The chi square test of independence for 2 x 2 contingency tables (1 *df*) was used on the appropriate table for the hypothesis being tested.

Table 1(A) indicates that a significantly higher proportion of father-absent boys than control boys made the crib choice ($p = .02$). There was no comparable difference in crib or bed choice between the girl groups. This finding supports Hypothesis 1 in which it was predicted that the father-absent boys would show more immaturity than the control boys, the crib being considered the more immature choice.

Table 1(B) shows that a significantly higher proportion of father-absent than control boys made more father-doll than mother-doll choices ($p = .05$). There was no comparable difference between the two groups of girls. This finding suggests support for Hypothesis 2 if it can be assumed that high father-doll choice implies strong strivings for father-identification.

Table 1(C) suggests support for Hypothesis 3. This table indicates that, from the mother interview, a larger proportion of father-absent boys showed high compensatory masculinity than control boys ($p = .01$).

Table 1(D) indicates that, from the mother-interview, a significantly higher proportion of father-absent than control boys showed poor peer adjustment ($p = .01$). The father-absent boys were also significantly more likely to show poor peer adjustment than father-absent girls ($p = .01$). This finding lends support to Hypothesis 4.

Table 1(E) shows that a significantly higher proportion of father-absent girls than control

girls chose the mother doll more often than the father doll in the Separation situation (p = .01). In fact, the father-absent girls differed significantly from all other groups, who did not differ significantly from each other in this regard. This finding will be discussed in connection with Hypothesis 5.

Table 1(F) suggests support for Hypothesis 5 concerning the direct effects of father-absence on dependency in girls. A significantly larger proportion of father-absent girls than control girls gave high dependency responses in the SDP (p = .05). There was no such difference for dependence in the SDP between the boy groups.

The measures on the draw-a-family task showed no significant differences between father-absent and control groups.

DISCUSSION

If it is correct that good peer adjustment is associated with adequate same-sex identification, then the fact that a higher proportion of control than father-absent boys showed adequate peer adjustment (Table 1[D]) might suggest that the father-absent boys were not as firmly identified with masculinity as were control boys. The fact that, as predicted (Hypothesis 4), a significantly higher proportion of father-absent boys than control boys or father-absent girls showed poor peer adjustment is consistent with the assumption that father-absence is more damaging to boys than girls.

Assuming, for the moment, that the father-absent boys were less securely identified with masculinity than were the control boys, then the lack of a significant difference between them with respect to frequency of high dependency responses in the SDP (Table 1[F]) may represent a compensatory masculine reluctance to express dependency. Despite the fact that the father-absent boys did not differ from the control boys regarding the measure of dependency, they did differ significantly in tending to choose the crib rather than the bed (Table 1[A]). We have suggested that this difference may represent underlying immaturity (Hypothesis 1).

If one assumes again that the father-absent boys were not as securely identified with their fathers as were control boys, then the results showing a significantly higher proportion of father-absent than control boys choosing the father more often than the mother doll (Table 1[B]) may imply greater strivings toward father-identification rather than its actual achievement (Hypothesis 2).

The father-absent girl, unlike the boy, still has the same-sex parent (mother) with her as a model for identification. The fact that a larger proportion of father-absent than control girls gave intense dependency responses in the SDP (Table 1[F]) may simply reflect actual increased dependency on the mother (Hypothesis 5). That this increased dependency was elicited in response to separation situations may indicate that, in the face of the virtual loss of one parent, the father-absent girl is particularly sensitive to fantasied separation threatening the loss of her model for identification (mother). This interpretation seems even more likely since a significantly higher proportion of the father-absent than control girls chose the mother doll as the parent they wished to return first in the Separation situation (Table 1[E]). It is interesting that the father-absent girls did not choose the crib rather than the bed significantly more often than the control girls. This suggests that, although more dependent, these girls, having their mothers with them as models for identification, continue to progress toward psychological maturity.

SUMMARY

This paper is concerned with the effects of father-absence on boys and on girls. In this study Norwegian families were compared with otherwise similar families in which the father was present.

The following hypotheses were made and generally supported by the findings:

1. More father-absent boys than father-present boys would show immaturity;

2. Being insecure in their identification with the father, father-absent boys would show stronger strivings toward father-identification than father-present boys;

3. The father-absent boys would react to their insecure masculine-identification with compensatory masculinity; and

4. The father-absent boys would demonstrate poorer peer adjustment than father-present boys and than father-absent girls.

5. It was also predicted that father-absent girls, perhaps threatened by the absence of one parent, would become more dependent on the mother than would father-present girls.

Father-absence appeared to be the determining variable in producing the significant results. However, this interpretation was advanced with caution since other possibly relevant independent variables could not be excluded.

REFERENCES

BACH, G. R. Father-fantasies and father-typing in father-separated children. *Child Developm.*, 1946, **17**, 63–80.

LYNN, D. B. Development and validation of a structured doll play test for children. *Quart. Bull., Indiana Univer. Med. Center*, 1955, Jan.

SEARS, R. R., PINTLER, MARGARET H., & SEARS, PAULINE S. Effect of father separation on preschool children's doll play aggression. *Child Develpm.*, 1946 **17**, 219–243.

TILLER, P. O. Father absence and personality development of children in sailor families. *Nord. Psykol. Monogr.*, 1958, Ser. No. 9.

Received July 18, 1958.

THE EFFECT OF AVERSIVE STIMULI ON REACTION TIME IN SCHIZOPHRENIA[1]

PETER J. LANG[2]
University of Buffalo

A NUMBER of theorists (Cameron, 1944; Hunt & Cofer, 1944; Wittman, 1937) have suggested that the impairment of schizophrenics on psychomotor, verbal, and concept tests is attributable to lowered "motivation," i.e., attitudinal variables such as indifference to the task or uncooperativeness. Attempts have been made to test this hypothesis by introducing aversive stimuli (shock, unpleasant "noise") into the experimental situation in ways designed to raise the motivational level of psychotics to perform. The present study was designed to examine this proposition for a disjunctive reaction time task and to clarify the role played by the aversive stimulus in modifying the schizophrenic's response.

Cohen (1956), Cavanagh (1958), and others (Pascal & Swensen, 1952; Rosenbaum, Grisell, & Mackavey, 1957a, 1957b) report that schizophrenics responding under escape conditions show a reduction in *psychological deficit* on tasks ranging from concept formation to simple reaction time. Their findings imply that the task becomes more "need relevant" in this context. The *S* overcomes his uncooperativeness or indifference to the task stimulus in an effort to escape the aversive stimulus (i.e., negative incentive). However, it is possible to explain the improvement of the schizophrenic *S*s, particularly on the reaction time tasks studied, in terms of a stimulus dynamism. The simultaneous presentation of the aversive stimulus and the discriminative stimulus may result in a greater "arousal" effect, in Hebb's (1956) terms, than is obtained by the latter stimulus alone. Differences in the improvement of normal and schizophrenic *S*s when aversive stimuli are introduced may reflect a difference in the response of normals and schizophrenics along the intensity dimension of the stimulus. The findings cited by Sands and Rodnick (1950) that schizophrenics show a "normal" GSR to an intense sound stimulus and a greatly diminished GSR to the word "ready" seems relevant to this point.

A second question raised by previous studies concerns the permanence of the observed improvement in performance. In none of the previous research have trials been administered subsequent to the withdrawal of the aversive stimulus. It is thus unclear whether the improvement represents permanent gain or is conditional on the presence of the newly introduced stimulus.

THE PROBLEM

Following a series of trials conducted under the usual conditions of administration for a reaction time (RT) task, *S*s responded while a collateral aversive stimulus ("white noise" at an unpleasant intensity) was presented through earphones. In addition to presenting the aversive stimulus in the escape context, it was introduced under avoidance and "excitation" conditions. In the latter case, the aversive stimulus was presented simultaneously with the discriminative stimulus, but its termination was not conditional on a correct response. Two control conditions were also included in the experiment. One group performed throughout the experiment under the usual conditions of administration for an RT task, and the second received information (nonaversive sound) following the pattern of the avoidance group. The design thus allowed for a separate analysis of the aversive stimulus as an incentive or "reinforcer" in the usual sense, and as a dynamic stimulus with inherent activating properties.

It was expected that if the aversive stimulus functioned primarily as an incentive, *S*s performing under escape and avoidance conditions would show significantly greater improvement

[1] This paper is adapted from a thesis submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Buffalo. The author wishes to express his appreciation to Walter Cohen, chairman of the doctoral committee, for his valuable counsel during the formulation and conduct of this research. Acknowledgement is also due Ira S. Cohen, Marvin Feldman, and James Drasgow for their many helpful suggestions.

[2] Now at the University of Pittsburgh.

than *Ss* performing under "excitation" or control conditions. A failure to find significant differences between the escape and excitation groups, on the other hand, would suggest that the improvement is a function of the "arousal" effect of an intense stimulus. In order to evaluate the extent to which improvement represented a persisting modification in performance, trials subsequent to the removal of the aversive stimulus were included in the design.

METHOD

Subjects

The *Ss* for this experiment were 70 male chronic schizophrenic patients and 50 normal males. The schizophrenic group was composed of 45 inpatients from Buffalo State Hospital, Buffalo, N. Y., and 25 inpatients from Gowanda State Homeopathic Hospital, Gowanda, N. Y. None displayed signs of neurological involvement or primary mental deficiency. The mean age was 35.04 years (σ = 4.82) and the modal educational level was 12 years (range: eight years to college graduate). The majority of the psychotics satisfied Drasgow's (1957) three-year statistical criterion of chronicity (mean number of years hospitalized was 4.65, (σ = 3.31), and all were considered to be chronically schizophrenic by the attending physician. None of these patients had received shock treatments during the three months previous to the experiment, and none were receiving barbiturates or tranquilizing drugs.

The normal *Ss* were selected from the surgical wards of the Veterans Administration Hospital, Buffalo, N. Y. They were seen either prior to surgery or while on convalescent status and were considered by the attending physician to be without residual effects from their illnesses that would impair their efficiency on the experimental task. They presented no history of psychiatric disorder; patients with physical complaints frequently termed "psychosomatic"—stomach ulcers, low back pains, asthma—were also excluded from the experiment. The mean age of these *Ss* was 30.08 years (σ = 7.64), and the modal educational level was 12 years (range: eighth grade to college graduate).

Apparatus

S was seated before a table on which a raised plywood board was placed. Attached to the end of the board, at an angle of 30 degrees to it, there was a gray plywood panel containing two neon lights horizontally arranged. Protruding from the base on either side of the panel were two standard telegraph keys. A rest button was affixed to the board directly in front of the *S*, 10.5 in. from both telegraph keys. *Ss* were instructed to execute a jump with the hand from the rest button to the right response key if the right light was illuminated, to the left response key if the left light was illuminated. Pressure on the correct response key terminated the light. A Standard Electric timer, activated simultaneously with the light and connected in series with the correct response key, recorded the "jump" interval to the nearest .01 sec.

In addition to the lights, a sound stimulus was introduced under some conditions of the experiment. The sound was provided by the output of a "white noise" generator, prerecorded on a one-hour tape at a constant level of 116 db. The sound was delivered to *S* through a pair of earphones which *S* wore throughout the experiment. The time of presentation and the duration of the sound stimulus was controlled by two Hunter Decade Interval Timers which were built into a wooden cabinet remote from the response keyboard. The cabinet also contained the recording clock, two ganged four-way mercury wall switches that activated the clock, the light and sound stimuli, and a double throw toggle switch which selected one or the other of the light-key circuits.

Procedure

All *Ss* were first introduced to the sound stimulus, regardless of later group assignment. Earphones were put on the *S* and the "noise" adjusted to an intensity reported as "really unpleasant, almost but not quite painful." Only one of the schizophrenic *Ss* became excessively disturbed by the "noise," and he was dropped from the experiment at this point. Both the normal and schizophrenic *Ss* showed little variability in the intensity selected. The level was approximately the same as that selected in response to the same instructions by a group of psychology graduate students.

All *Ss* were then given seven trials at the apparatus which were not used in the analysis. Subsequent to these warm-up trials, all *Ss* received 63 test trials divided into three series of 21 trials each—prestimulation, stimulation, and poststimulation. Within any series, trials were administered in blocks of seven with a 45-sec. rest between each block. Between each series of 21 trials, *Ss* rested for 1.5 min.

A trial consisted of one jump reaction to either the left or right key at the illumination of the appropriate light. A verbal ready signal was presented randomly between two and five sec. before the light. The sequence of left and right lights was determined from a table of random numbers, with the restriction that neither the left nor right light would be presented more than three times in succession.

Both schizophrenic and nonpsychiatric patients were divided into five groups—Control, Escape, Avoidance, Excitation, and Information. *Ss* were assigned to these groups randomly within blocks of five from each diagnostic classification. Thus, from every five *Ss* selected, one member was assigned to each of the five experimental groups. Fourteen schizophrenics and ten normal *Ss* composed each group.

The prestimulation trials were conducted under the usual conditions of administration for an RT task. All *Ss* were simply instructed to respond as quickly as they were able. However, depending on group assignment, *Ss* performed during stimulation trials under one of the following conditions:

1. Control (CO). The "noise" was not presented.
2. Escape (ES). The "noise" was turned on simul-

taneously with the light and terminated by a correct response.

3. Excitation (EX). The "noise" was turned on simultaneously with the light and maintained for a fixed duration. For each block of seven trials, the sound was presented twice at both .40 sec. and 1.20 sec., and three times at .80 sec. The sequence of durations was random.
4. Avoidance (AV). The "noise" was presented after a time interval subsequent to the light and was maintained for 1.20 sec. However, depression of the correct key before the end of the time interval prevented the "noise" from coming on. The time interval from the onset of the light to the onset of the sound was determined by each *S*'s median score on the previous block of seven trials.
5. Information (INF). The "noise" was presented at low intensity subsequent to the visual stimulus. The pattern of presentation was the same as that described for the AV condition.

Prior to the stimulation trials, *Ss* were instructed appropriately, depending on their condition assignment. The ES group was told that their response would terminate the "noise." *Ss* performing under EX conditions were instructed that their response was unrelated to the termination of the "noise." The AV and INF groups were told that the sound would be presented subsequent to their response if they were "too slow."

The experimental groups composed of psychotic *Ss* are hereafter labeled PCO, PES, PEX, etc. Groups composed of normal *Ss* are labeled NCO, NES, etc.

Following the stimulation trials, 21 poststimulation trials were administered in which no sound was presented to any group. Instructions were the same for all groups and indicated only that *S* should respond to the appropriate key as quickly as he could. Questions about the absence of the sound were answered by reinstructing *S* to respond as soon as he saw one of the lights illuminated.

RESULTS

Ss' median scores for each block of seven reaction times were calculated. Thus, for each series of trials (prestimulation, stimulation, and poststimulation) three median scores were available. The score selected to represent the *S*'s performance for any series was the fastest median RT for that series. This score was chosen because pilot data revealed it to be a reasonably stable indication of *S*'s best effort at each stage of performance.

Differences between schizophrenic and normal *Ss* were first analyzed in terms of the mean score for all *Ss* in a diagnostic group on the prestimulation series. A *t* test of the significance of the difference between means and an *F* test for homogeneity of variance were performed. Table 1 reveals that both these statistics were significant well beyond the .01 level.

TABLE 1
PERFORMANCE OF SCHIZOPHRENIC AND NORMAL GROUPS ON PRESTIMULATION TRIALS

Diagnostic Group	Mean RT[a]	*SD*	*t*	*F*
Normal	40.02	4.46	7.31*	22.14*
Schizophrenic	62.09	21.01		

[a] Statistics are reported in units of .01 sec.
* Significant at .01 level.

To determine whether the differences in means were significant irrespective of variance, an approximate test of this assumption described by Edwards (1950) was employed. The initially obtained *t* (7.31) greatly exceeded the approximate *t* (2.54) necessary for significance at the .01 level. These results confirm the findings of others (Huston & Shakow, 1937; King, 1954; Stotsky, 1957) that under the usual conditions of administration, schizophrenic *Ss* display clear evidence of psychological deficit on an RT task.

The performances of the schizophrenic and normal experimental groups (including CO condition) were separately analyzed. In neither case was there a significant difference between any of the experimental groups during prestimulation trials. The analysis of variance of the normal experimental groups yielded an *F* of 1.12. A similar analysis yielded an *F* of .22 for the schizophrenic experimental groups. because of the nonnormal character of the data obtained from the schizophrenic *Ss*, a logarithmic transformation was imposed prior to the variance analysis.

The stimulation and poststimulation trials were analyzed in terms of decrement scores. Because these data were not normally distributed, nonparametric statistics were employed in subsequent analyses. Decrement scores were obtained by separately subtracting the stimulation and poststimulation scores of each *S* from his prestimulation trials score. This is the method used by Stotsky (1957) in a previous study of the effect of "motivating" conditions on the RT of schizophrenic *Ss*.

TABLE 2
KRUSKAL-WALLIS ANALYSES OF VARIANCE OF THE SCHIZOPHRENIC AND NORMAL GROUPS

Diagnostic Group	Trials	*H* (corrected for ties)
Normal	Stimulation	9.18
	Poststimulation	7.99
Schizophrenic	Stimulation	15.42**
	Poststimulation	11.79*

* Significant at .02 level.
** Significant at .01 level.

Simple subtraction rather than ratio scores were employed because the experimental groups displayed no consistent relationship between level of performance and decrement. The mean of the rank order correlations between prestimulation score and simple decrement from prestimulation to stimulation trials for the five schizophrenic experimental groups was +.06 (range: +.45 to −.26). None of the correlations was significant.

Normal Ss

Two Kruskal-Wallis analyses of variance were performed (Siegel, 1956). In the first analysis, decrement scores for stimulation trials were examined; in the second, scores for the poststimulation series were evaluated. Neither analysis of variance was significant (Table 2), suggesting that the conditions of information and aversive stimulation used in this study do not significantly alter the disjunctive reaction time of normal *Ss*.

Schizophrenic Ss

The same technique used in analyzing the decrement scores of the normal *Ss* was applied to data from the schizophrenic patients. The results are reported in Table 2. The variance analyses of the stimulation and the poststimulation series were both significant at the .01 and .02 levels respectively.

The Wilcoxon test for unpaired replicates (Wilcoxon, 1949) was used to evaluate differences between specific groups. The PES, PEX, and PINF groups all proved to be significantly different from the PCO group for the stimulation series.

Except for the PINF group, none of the experimental groups displayed decrement scores significantly different from those of CO *Ss* on poststimulation trials. The results of all significance tests between psychotic groups are reported in Table 3.

TABLE 3
WILCOXON TESTS OF DIFFERENCES IN DECREMENT SCORES BETWEEN THE SCHIZOPHRENIC EXPERIMENTAL GROUPS

Schizophrenic Groups	PCO	PES	PEX	PINF
Stimulation Series Trials				
PES	128.5**			
PEX	135.5**	189.0		
PINF	142.5**	180.0	195.0	
PAV	162.0	166.5	178.0	185.0
Poststimulation Series Trials				
PES	183.5			
PEX	181.0	171.0		
PINF	158.0*	142.0**	177.0	
PAV	163.0	146.0**	177.5	195.0

Note.—Table entries are the least sum of ranks obtained for each Wilcoxon test. For an *N* of 14 a rank total of 160 or less indicates a significant difference between groups at the .05 level of confidence.
* Significant at .05 level.
** Significant at .01 level.

Comparison of Diagnostic Groups

Regardless of the conditions of administration, schizophrenic patients displayed longer reaction time latency than did normal *Ss*. A Mann-Whitney *U* test of the difference between the normal experimental group with the slowest median reaction time for any test series (NES, prestimulation) and the schizophrenic experimental group with the fastest median time (PEX, stimulation) yielded a *U* of 25.0 ($p < .02$).

The Mann-Whitney test was also employed to evaluate differences between the normal and schizophrenic *Ss* under each experimental condition. Schizophrenic *Ss* seemed slower to profit from simple practice than normal *Ss*. A stimulation-trials comparison of the decrement scores of the normal and schizophrenic *Ss* under CO conditions yielded a *U* of 30.05 ($p < .05$). The decrement scores for these *Ss* and for those of the other experimental groups are presented in Table 4.

With the exception of the difference between the PEX and NEX group, none of the other differences between the decrement scores of normal and schizophrenic *Ss* proved to be

TABLE 4
PERFORMANCE OF THE NORMAL AND SCHIZOPHRENIC EXPERIMENTAL GROUPS

Experimental Condition	Median Decrement Score	
	Stimulation Trials	Poststimulation Trials
NCO	3.5	2.5
NES	4.0	1.5
NEX	1.0	0.5
NAV	2.0	1.5
NINF	2.5	3.5
PCO	1.0	1.0
PES	6.0	0.0
PEX	6.0	2.5
PAV	3.0	3.0
PINF	5.0	5.0

Note.—All scores reported in units of .01 secs.

significant for either the stimulation or poststimulation trials. The evaluation of the PEX and NEX groups yielded a U of 32.0 ($p < .05$) for the stimulation series.

DISCUSSION

The results of the present research confirmed the findings of others (Cohen, 1956; Pascal & Swensen, 1952; Rosenbaum et al., 1957a) that under conditions of aversive stimulation (escape), schizophrenic *S*s show a greater decrement in reaction time than schizophrenics performing under the usual conditions of administration. However, the suggestion that improvement in performance under escape conditions may not be a function of the reinforcing effect of termination of the aversive stimulus was borne out. Both the PES and PEX groups proved to be significantly different from PCO *S*s during stimulation trials and displayed no mutual differences in decrement scores. Furthermore, the PAV group was not successfully differentiated from either the PCO or PINF groups. Thus, the aversive stimuli did not seem to function in ways which made a faster response more important to the schizophrenic. Improvement under these conditions can be accounted for simply in terms of the stimulus dynamism.

Furthermore, the improvement of chronic schizophrenics was not marked when compared with the degree of deficit. This finding is not inconsistent with results obtained by others (Cohen, 1956; Peters, 1953; Rosenbaum et al., 1957a; Stotsky, 1957) using either positive or negative reinforcers and seems to call for one or two conclusions: Either the conditions so far used have not been appropriate for motivating chronic schizophrenics, or lack of motivation is not the major factor in reaction time deficit. The significant improvement of the PINF group[3] in the present study seems to offer some evidence for the latter view. As McGeogh and Irion (1952, p. 266) pointed out, "the learner must want to make the correct response before the information signals of Right and Wrong can have much efficacy." Perhaps schizophrenics who are willing to take part in experiments, are already reasonably well "motivated" to perform.

Nevertheless, the fact that schizophrenics were able to improve their performance would suggest that response latency is not circumscribed by some organic defect of the peripheral nervous system. The failure of chronic schizophrenics to perform at the level of normal *S*s seems related to a fundamental inability to make adequate use of stimuli in the usual experimental setting. Shakow's concept (1950) of an inability to maintain set appears to fit the data, both in terms of readiness to respond to a specific stimulus and "readiness to respond to a generalization from a group of stimuli."

The factors which underly this impaired set are not revealed in this experiment. Some defect in hypothesized central facilitating systems of the lower brain, such as suggested by Heath and his co-workers (1954), may be involved. On another conceptual level, special preoccupations, like idiosyncratic thought processes which interfere with attention, may be significant. However, the present data

[3] These *S*s improved as much as any other group during stimulation trials. Furthermore, this was the only group which learned a new response to the visual stimulus, having maintained the improvement during the poststimulation series. This result appears to be incompatible with the thesis of Rodnick and Garmezy (1957). These latter writers hold that "social censure," which amounts to information about errors in the studies they report, impairs the performance of schizophrenic *S*s rather than improves it. However, the inconsistency may result from differences in method, particularly the degree to which psychological stress is associated with the presentation of the information. A further study is called for to determine whether the information provided in this experiment, delivered in a more punitive fashion, would produce the same result reported here.

would suggest that the interference is not readily modified by artificial efforts to manipulate the *S*'s "need" to overcome it.

SUMMARY

The purpose of the experiment was to assess the motivational role of aversive stimuli in modifying the performance of normal and psychotic *S*s. Seventy schizophrenics and 50 normal males performed a visual reaction time task in three series of 21 trials each. During the second series *S*s responded under one of the following experimental conditions—Control, Information, Avoidance, Excitation, and Escape. The latter three conditions involved the use of intense noise as an aversive stimulus.

None of the experimental conditions appreciably altered the performance of normal *S*s. However, schizophrenic *S*s performing under Escape, Excitation, and Information conditions displayed greater improvement than control *S*s.

The results were interpreted as suggesting that the "need relevance" of the task is not a major factor in the reaction time deficit of chronic schizophrenics. The improvement of schizophrenic *S*s under aversive stimulus conditions was explained in terms of the dynamogenic properties of an intense stimulus.

REFERENCES

CAMERON, N. The functional psychoses. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944. Pp. 861–921.

CAVANAUGH, D. Improvement in the performance of schizophrenics on concept formation tasks as a function of motivational change. *J. abnorm. soc. Psychol.*, 1958, **57**, 8–12.

COHEN, B. C. Motivation and performance in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **52**, 186–190.

DRASGOW, J. A criterion for chronicity in schizophrenia. *Psychiat. Quart.*, 1957, **31**, 454–457.

EDWARDS, A. *Experimental design in psychological research*. New York: Rinehart, 1950.

HEATH, R. G. (Ed.) *Studies in schizophrenia*. Cambridge, Mass.: Harvard Univer. Press, 1954.

HEBB, D. O. Drives and the C.N.S. *Psychol. Rev.*, 1956, **62**, 243–254.

HUNT, J. McV., & COFER, C. N. Psychological deficit. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944. Pp. 971–1032.

HUSTON, P. E., & SHAKOW, D. Studies of motor functioning in schizophrenia: II. Reaction time. *J. genet. Psychol.*, 1937, **16**, 39–82.

KING, H. E. *Psychomotor aspects of mental disease*. Cambridge, Mass.: Harvard Univer. Press, 1954.

MCGEOCH, J., & IRION, A. *The psychology of human learning*. (2nd ed.). New York: Longmans, Green, 1952.

PASCAL, G. R., & SWENSEN, C. Learning in mentally ill patients under conditions of unusual motivation. *J. Pers.*, 1952, **21**, 240–249.

PETERS, H. N. Multiple choice learning in the chronic schizophrenic. *J. clin. Psychol.*, 1953, **9**, 328–333.

RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. Jones (Ed.), *Nebraska symposium on motivation*. Vol. 5. Lincoln, Neb.: Univer. Nebraska Press, 1957. Pp. 109–184.

ROSENBAUM, G., GRISELL, J. L., & MACKAVEY, W. R. Effects of biological and social motivation on schizophrenic reaction time. *J. abnorm. soc. Psychol.*, 1957, **54**, 364–368. (a)

ROSENBAUM, G., GRISELL, J. L., & MACKAVEY, W. R. The relationship of age and privilege status to reaction time indices of schizophrenic motivation. *J. abnorm. soc. Psychol.*, 1957, **55**, 202–207. (b)

SANDS, S. L., & RODNICK, E. H. Concept and experimental design in the study of stress and personality. *Amer. J. Psychiat.*, 1950, **106**, 673–679.

SHAKOW, D. Some psychological features of schizophrenia. In M. L. Reymert (Ed.), *Feelings and emotions*. New York: McGraw-Hill, 1950. Pp. 383–390.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

STOTSKY, B. A. Motivation and task complexity as factors in the psychomotor responses of schizophrenics. *J. Pers.*, 1957, **25**, 327–343.

WILCOXON, F. *Some rapid approximate statistical procedures*. (rev. ed.) New York: American Cyanamid, 1949.

WITTMAN, PHYLLIS. An evaluation of opposed theories concerning the etiology of so-called "dementia" in dementia praecox. *Amer. J. Psychiat.*, 1937, **93**, 1363–1372.

Received July 16, 1958.

CASE REPORT

THE ELIMINATION OF TANTRUM BEHAVIOR BY EXTINCTION PROCEDURES

CARL D. WILLIAMS

University of Miami

This paper reports the successful treatment of tyrant-like tantrum behavior in a male child by the removal of reinforcement. The subject (*S*) was approximately 21 months old. He had been seriously ill much of the first 18 months of his life. His health then improved considerably, and he gained weight and vigor.

S now demanded the special care and attention that had been given him over the many critical months. He enforced some of his wishes, especially at bedtime, by unleashing tantrum behavior to control the actions of his parents.

The parents and an aunt took turns in putting him to bed both at night and for *S*'s afternoon nap. If the parent left the bedroom after putting *S* in his bed, *S* would scream and fuss until the parent returned to the room. As a result, the parent was unable to leave the bedroom until after *S* went to sleep. If the parent began to read while in the bedroom, *S* would cry until the reading material was put down. The parents felt that *S* enjoyed his control over them and that he fought off going to sleep as long as he could. In any event, a parent was spending from one-half to two hours each bedtime just waiting in the bedroom until *S* went to sleep.

Following medical reassurance regarding *S*'s physical condition, it was decided to remove the reinforcement of this tyrant-like tantrum behavior. Consistent with the learning principle that, in general, behavior that is not reinforced will be extinguished, a parent or the aunt put *S* to bed in a leisurely and relaxed fashion. After bedtime pleasantries, the parent left the bedroom and closed the door. *S* screamed and raged, but the parent did not re-enter the room. The duration of screaming and crying was obtained from the time the door was closed.

The results are shown in Fig. 1. It can be seen that *S* continued screaming for 45 min. the first time he was put to bed in the first extinction series. *S* did not cry at all the second time he was put to bed. This is perhaps attributable to his fatigue from the crying of Occasion 1. By the tenth occasion, *S* no longer whimpered, fussed, or cried when the parent left the room. Rather, he smiled as they left. The parents felt that he made happy sounds until he dropped off to sleep.

About a week later, *S* screamed and fussed after the aunt put him to bed, probably reflecting spontaneous recovery of the tantrum behavior. The aunt then reinforced the tantrum behavior by returning to *S*'s bedroom and remaining there until he went to sleep. It was then necessary to extinguish this behavior a second time.

Figure 1 shows that the second extinction curve is similar to the first. Both curves are generally similar to extinction curves obtained with sub-

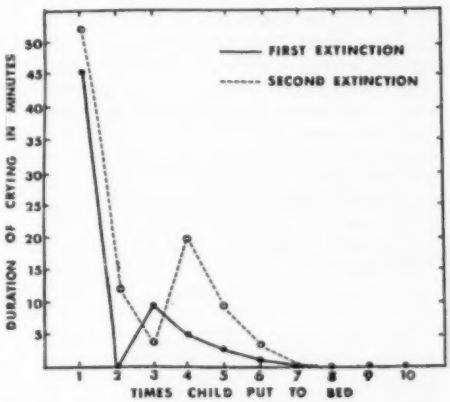

FIG. 1. LENGTH OF CRYING IN TWO EXTINCTION SERIES AS A FUNCTION OF SUCCESSIVE OCCASIONS OF BEING PUT TO BED

human subjects. The second extinction series reached zero by the ninth occasion. No further tantrums at bedtime were reported during the next two years.

It should be emphasized that the treatment in this case did not involve aversive punishment. All that was done was to remove the reinforcement. Extinction of the tyrant-like tantrum behavior then occurred.

No unfortunate side- or aftereffects of this treatment were observed. At three and three-quarters years of age, *S* appeared to be a friendly, expressive, outgoing child.

Received October 1, 1958.

CRITIQUE AND NOTES

SOCIAL CLASS AND THE NATURE OF THE INCENTIVE IN DISCRIMINATION LEARNING[1]

GLENN TERRELL, JR., KATHRYN DURKIN
University of Colorado
AND MELVIN WIESLEY
Fairview School, Boulder, Colorado

PREVIOUS research has suggested interesting differences in the relative effectiveness of material and nonmaterial incentives for children of different social-class backgrounds. In an experiment involving the solution of a series of tasks, Douvan (1956) found that middle-class subjects (*Ss*) maintained approximately the same level of achievement motivation when told they had reached a norm as when they were promised a sum of money. The motivation of lower-class *Ss*, on the other hand, dropped significantly when the material reward was absent. Terrell and Kennedy (1957) found that children, a preponderance of whom were from a rural background, require significantly more trials to learn a "larger-than" response when given only a light flash as an indication of a correct response than when given a series of material incentives, including candy, praise, token, and reproof. In a similar experiment performed with middle-class children, Terrell (1958) found that children assigned to the light flash condition learned somewhat faster than those who received candy.

From the foregoing, it would be predicted that an interaction exists between social class and the nature of the incentive. Specifically, it was believed that a nonmaterial incentive is as effective as a material incentive for middle-class *Ss*, whereas, for lower-class *Ss* a material incentive is more effective than a nonmaterial one. The present experiment was designed to test this belief. It is important to note that while Douvan studied the relationship between these variables using a measure of motivation as a criterion variable, this paper is a report of a discrimination learning experiment, involving an acquisition measure as a dependent variable.

METHOD

Subjects. There were 12 *Ss* in each of the following age categories: 5-, 6-, 10-, and 11-year-olds, with an equal number of boys and girls in each age group. The school from which *Ss* for this experiment were drawn very appropriately contains children from a wide variety of social-class backgrounds. Warner's Index of Status Characteristics (Warner, Meeker, & Eels, 1949) was the measure used to define class position. In the present study, all *Ss* designated as middle class fell into either the upper-middle or the lower-middle classes of Warner's schema, while *Ss* with scores placing them in Warner's upper-lower and lower-lower classes constituted the lower class.

Materials. The apparatus for the problem has been described in detail elsewhere (Terrell, 1957). There were three pairs of three-dimensional geometric figures in the shape of cubes, cones, and cylinders. The small member of each stimulus set had a basal area of 4 sq. in., while the large member had a basal area of 8 sq. in. These stimuli are hereafter referred to as the training stimulus sets. A third cube with a basal area of 16 sq. in. was used in a transposition test, along with the 8 sq. in. cube. These stimuli are hereafter referred to as the test stimuli. The order of presenting the stimuli and the position of the positive, large-size, stimulus were randomized alike for each *S*.

Additional apparatus consisted of a 16 x 24 x 4 in. box which contained the batteries and circuits necessary to operate a signal light. Two jacks and two push-button mounts were on top of the box. The stimuli were placed into the jacks on each trial. Locked onto the rear edge of the box was a 10 x 16 x ¼ in. panel board which contained the signal light. The circuits were arranged so that a correct response, pushing the button at the base of the large stimulus, caused the light to go on.

Design. There were two experimental groups. Following each correct response, one group of *Ss* received a nonmaterial reward, a light flash, while the *Ss* of the other group received a material reward, a small piece of candy in addition to the light flash. Within each of the two levels of social class, *Ss* were randomly assigned to the two incentive conditions, making a total of 12 *Ss* in each treatment-level combination.

Procedure. The *Ss* were tested individually. Each *S* received the following instructions:

[1] The writers wish to thank Nevin J. Platt and all of the teachers of Fairview school, without whose cooperation this experiment could not have been performed.

This is a game where I want you to try to choose one of these (*E* points to the training stimulus sets), and push the button in front of the one you choose. If you are right, this little light will go on. If you are wrong, the light will not go on. Now remember, the game is to see how quickly you can learn to choose the one that makes the light go on.

The last sentence was repeated after every tenth trial. Immediately after reaching a criterion of 9 of 10 correct responses, each *S* was given a four-trial transposition test on the test stimuli. The same differential incentive conditions employed during the acquisition trials were continued during the test trials.

RESULTS AND DISCUSSION

The original design called for an analysis of variance of the mean number of trials to the criterion and mean number of transposition responses. All *S*s except one transposed on every trial, making an analysis of transposition data useless. In the case of the number of trials to the criterion, the homogeneity and normality assumptions were far from met. Also, the criterial means and variances were correlated. For these reasons, it seemed advisable to subject these data to a nonparametric test. Wilson's (1956) chi-square technique makes possible the test of hypotheses concerning main and interaction effects ordinarily tested by a two-way analysis of variance.

Table 1 contains the training means and *SD*s. As can be seen from this table, a striking interaction exists between social class and the type of incentive. Middle-class children learn more quickly when given a nonmaterial incentive than when given a material incentive, while the reverse is true of lower-class children. The over-all means for social class are identical, while the difference between over-all incentive means is negligible.

Table 2 contains frequencies above and below the median for each of the incentive-class combinations in number of trials to the criterion. All analyses and subsequent discussion are based on the data of Table 2. The predicted interaction is apparent in this table, and to a significant degree. The chi-square value for interaction is 5.17, which with one degree of freedom reaches significance at about the .03 significance level.[2] The direction of the interaction is the same for ages 5 and 6 as well as for ages 10 and 11. No analyses of interaction effects were made for these ages independently, since the frequencies were very small. The chi-square test for main effects of class was nonsignificant, while the test of main effects of incentive was significant at the .05 level. There

[2] Yates' correction was applied to the computation of the main effects of social class and incentive conditions.

TABLE 1

MEANS AND *SD*s OF TRIALS TO CRITERION IN TRAINING (Each treatment group $N = 20$)

Class	Material Incentive: Present		Material Incentive: Absent		Total	
	Mean	*SD*	*M*	*SD*	*M*	*SD*
Middle	7.91	8.33	3.50	2.58	5.70	6.72
Lower	3.41	3.02	8.00	5.15	5.70	4.80
Total	5.66	6.66	5.75	4.65		

TABLE 2

FREQUENCIES ABOVE AND BELOW THE MEDIAN IN NUMBER OF TRIALS TO CRITERION FOR EACH INCENTIVE-CLASS COMBINATION

	Incentive: Middle Class		Incentive: Lower Class	
	Material	Nonmaterial	Material	Nonmaterial
Above Median	6	7	5	12
Below Median	6	5	7	0

was a significant tendency for *S*s assigned to the material incentive condition to learn more quickly than those *S*s assigned to the nonmaterial incentive treatment. This finding is rendered rather meaningless in view of the significant interaction present in this experiment.

Several interesting implications arise from this experiment. There is evidence to indicate that parents of middle-class children place a greater emphasis on learning for learning's sake than do parents of lower-class children (Davis, 1944; Erickson, 1947). Additional evidence in support of this hypothesis was found in the aforementioned experiment by Terrell (1958), in which middle-class *S*s indicated in a questionnaire given to them following the experiment that they would rather "do something for the fun of it" than to be "promised or given something for doing it." It would appear that the most important feature in the learning of middle-class *S*s is merely some indication that they are progressing. It is strikingly apparent in Table 2, however, that the presence of a material incentive is very important to lower-class *S*s. It is possible that the lower-class child is too preoccupied with obtaining the material, day-to-day necessities of life to have the opportunity to learn the value of less material, symbolic incentives. Davis (1941, 1943) suggests such a possibility. Additionally, since it is likely that the lower-class child generally is more deprived of the specific material incentive, candy, used in the

present study, this deprivation may result in an intensification of his desire for candy. It would be interesting to know whether or not the same results would be obtained in experiments employing other material incentives such as toys, clothing, movie tickets, and the like.

Finally, there is a possibility that the middle-class child is more adept than the lower-class child at engaging in effective imaginative activity during learning.[3] If this be the case, it would seem that the middle-class child would learn more effectively under a symbolic, nonmaterial type incentive than lower-class children. Research bearing on this point is currently being conducted by the senior author.

[3] The writers are indebted to Howard E. Gruber for this interesting suggestion.

REFERENCES

DAVIS, A. American status systems and the socialization of the child. *Amer. sociol. Rev.*, 1941, **6**, 345–354.

DAVIS, A. Child training and social class. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds), *Child behavior and development*. New York: McGraw-Hill, 1943.

DAVIS, A. Socialization and adolescent personality. *Adolescence, Forty-third Yearbook, Part I*. Chicago: National Soc. for Study of Education, 1944.

DOUVAN, E. Social status and success striving. *J. abnorm. soc. Psychol.*, 1956, **52**, 219–223.

ERICKSON, MARTHA C. Social status and child rearing practices. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology*. New York: Holt, 1947. Pp. 494–501.

TERRELL, G. The role of incentive in discrimination learning in children. *Child Develpm.*, 1958, **29**, 231–236.

TERRELL, G., & KENNEDY, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, **53**, 257–260.

WARNER, W. L., MEEKER, M., & EELS, K., *Social class in America*. Chicago: Science Research Associates, 1949.

WILSON, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, **53**, 96–101.

Received May 5, 1958.

INTELLECTUAL AND PERSONALITY CORRELATES OF TEST ANXIETY[1]

IRWIN G. SARASON[2]
University of Washington

THE present report is concerned with providing empirical evidence concerning interrelationships among different measures of anxiety and the relationships of these anxiety measures to intellectual performance. The study was designed so as to shed light on the following questions: (*a*) What is the relationship of a specific anxiety—test anxiety—to more general measures of anxiety? (*b*) Is the measurement of test anxiety as confounded with social desirability as more general anxiety indices appear to be? (*c*) Does test anxiety differ from measures of general anxiety and social desirability in the extent to which it is related to several measures of intellectual performance?

If test anxiety is to be used as a separate construct partially independent of general anxiety, it would seem important at least to demonstrate that a measure of test anxiety correlates with general anxiety measures to only a moderate extent and also that test anxiety is a superior predictor of performance in test situations when compared with general indices of anxiety.

[1] This research was supported by grants from the Agnes H. Anderson Research Fund and the Graduate School Research Fund of the University of Washington. Portions of this paper were presented at the April, 1958, meeting of the Western Psychological Association.

[2] The writer is indebted to Calvin E. Wright for his aid in computing the statistical analysis of the data. The cooperation of R. B. Loucks, B. B. McKeever, and William J. Meads in making subjects available is greatly appreciated.

METHOD

Subjects

The *Ss* were 309 freshman and sophomore male students enrolled in introductory psychology and Army ROTC courses at the University of Washington and 67 freshman and sophomore female students enrolled in introductory psychology courses at the same institution.

Materials

The following personality tests were administered to the *Ss* during class meetings: (*a*) Sarason's (1958) true-false Test Anxiety (TA) scale derived from items in the Test Anxiety Questionnaire used by Mandler and Sarason (1952); (*b*) the Lack of Protection scale (LP), items of which are purported to be a general measure of separation anxiety (Sarason, 1958); (*c*) Bendig's (1956) 20-

TABLE 1
INTERCORRELATIONS OF PERSONALITY AND INTELLECTUAL MEASURES FOR 309 MEN

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
TA	1		−23	41	36	−12	−03	−14	−12	−12	−01	−27	−07	−20	−11	−19	−24	−09
SDS	2			−50	−44	06	−03	−04	03	00	08	−07	−07	08	−05	−05	03	−03
MAS	3				55	−08	−08	−03	−09	−08	−01	01	−04	−12	01	−09	−04	−03
LP	4					−03	03	00	−01	04	−01	05	−02	−06	04	−04	02	02
High School Eng. average	5						58	68	73	66	50	31	−11	53	40	26	30	24
High School Math. average	6							60	50	64	42	13	17	38	25	58	18	41
Foreign Lang. average	7								61	64	43	31	00	46	36	38	18	27
Soc. Sci. average	8									66	44	27	−05	36	20	21	32	20
Natural Sci. average	9										46	31	15	36	14	43	24	24
Elective average	10											02	−03	26	17	19	04	19
G-Z Verbal	11												05	51	49	35	62	28
G-Z Mech. Knowledge	12													−05	−09	46	18	22
Coop. Eng. Usage	13														62	44	46	46
Coop. Eng. Spelling	14															27	32	33
Coop. Math. I	15																41	57
Coop. Soc. Stud. II	16																	40
ACE L score (1948)	17																	

r = .113 at .05 level.
r = .148 at .01 level.

item short form of the MAS; (*d*) Edwards's (1957) 39-item Social Desirability Scale (SDS).

In addition, each entering student at the University of Washington is routinely administered a battery of intellectual performance tests. These tests, plus several available high school grade-point averages, were also used in the present study.[3] The following are the intellectual measures which were available: (*a*) High school English grade-point average (GPA); (*b*) High school mathematics GPA; (*c*) High school foreign language GPA; (*d*) High school social science GPA; (*e*) High school natural science GPA; (*f*) High school elective GPA; (*g*) Verbal score on the Guilford-Zimmerman Aptitude Survey; (*h*) Mechanical knowledge score on the Guilford-Zimmerman Aptitude Survey; (*i*) Cooperative English Usage score; (*h*) Cooperative English Spelling score; (*k*) Cooperative Mathematics I score; (*l*) Cooperative Social Studies II score; (*m*) 1948 ACE L scores.

Obviously, one limitation on the generality of the results obtained was the fact that the personality and intellectual measures were not obtained contemporaneously.

Procedure

From the intellectual and personality measures, intercorrelation matrices were computed separately for the males and females. The total number of these measures was 17; consequently, the results were in the form of two 17 × 17 matrices.

[3] The writer is indebted to Paul Horst for his cooperation in making available these data. The writer is also indebted to Bernard H. Taylor for his assistance in gathering this data.

RESULTS

The two 17 × 17 matrices are presented in Tables 1 and 2 for the males and females respectively. With respect to the question concerning the relationship of test anxiety to more general anxiety indices, it would appear that TA, for both men and women, correlates to a very much lower extent with MAS and LP than these scales correlate with each other or than their reported reliabilities (Bendig, 1956; Sarason, 1958). Clearly, the correlations between TA and MAS of .41 and .49 for men and women, respectively, and the correlations of .36 and .44 between TA and LP for men and women, respectively, indicate that there could only be a relatively small proportion of the TA variance accountable by its correlation with the two more general indices.

Another question was whether or not TA differed from MAS and LP in the extent to which *S*s display a tendency to respond on the basis of the social desirability of test items as measured by the SDS. For both men and women, TA correlated significantly less negatively with the SDS than did either MAS or LP.

It is interesting to note that when the intercorrelations of the four personality tests are considered, in every case the tests are more highly intercorrelated for the females than for the males.

The final question concerned the relationships of the four personality measures to the intellectual

TABLE 2
INTERCORRELATIONS OF PERSONALITY AND INTELLECTUAL MEASURES FOR 67 WOMEN

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
TA	1		−49	49	44	−25	−18	−09	−29	−26	−12	−16	−03	−37	−20	−23	−29	−36
SDS	2			−79	−68	15	−14	00	19	09	07	−18	04	24	13	−04	04	23
MAS	3				60	−14	03	−04	−12	−15	05	21	−07	−17	05	05	−05	−21
LP	4					−12	05	−09	−09	−11	−16	−01	−14	−19	−01	−08	−05	−11
High School Eng. average	5						63	62	70	60	57	47	−04	64	40	47	45	35
High School Math. average	6							60	57	61	38	42	16	50	27	58	41	53
Foreign Lang. average	7								51	65	41	41	−03	47	29	40	34	28
Soc. Sci. average	8									52	41	27	01	44	42	36	38	28
Natural Sci. average	9										42	29	10	42	22	46	19	25
Elective average	10											18	09	44	30	31	12	23
G-Z Verbal	11												−04	47	28	43	61	33
G-Z Mech. Knowledge	12													−03	01	39	−05	25
Coop. Eng. Usage	13														40	50	59	62
Coop. Eng. Spelling	14															23	26	25
Coop. Math. I	15																53	66
Coop. Soc. Stud. II	16																	57
ACE L score (1948)	17																	

r = .250 at .05 level.
r = .325 at .05 level.

measures. Each of the four personality measures was correlated with the 13 intellectual measures. The results quite conclusively point to a greater degree of relationship between TA and these intellectual indices than between MAS, LP, or SDS and these estimates of intellect. For the SDS and LP scales, there were no significant correlations with any of the 13 intellectual measures for either men or women. The MAS scale correlated significantly with one of the 13 intellectual measures for men. No significant correlations were obtained for women.

Contrasted with these null findings are the results for the TA scale. For the men, TA correlated significantly with 8 of the 13 intellectual variables; for the women, the number of significant correlations was six. In every case these correlations were negative. Inspection of Tables 1 and 2 reveals that variables which correlate significantly for men do not always correlate significantly for women. However, both men and women showed significant negative correlations between TA and such intellectual measures as high school English, social sciences and natural science averages, Cooperative English Usage, and Cooperative Social Studies II tests.

DISCUSSION

Although TA did correlate significantly with MAS and LP, the results suggest the hypothesis that common variance between TA and the general anxiety indices is sufficiently small to make worthwhile further independent study of test anxiety. The correlations of TA with the general anxiety measures of MAS and LP are quite in line with the previous finding by Gordon and Sarason (1955) of only a moderate, although significant, correlation (r = .468) between graphic rating scales of test anxiety and general anxiety. Sarason (1958) similarly found a correlation of .46 between true–false test and general anxiety scales for *Ss* drawn from a neuropsychiatric patient population.

Rather noteworthy in the TA results were the relatively low TA–SD correlations for both men and women. At this point, it is not clear why admitting to anxiety in testing situations should be less confounded by social desirability than admitting to anxiety in a wide variety of situations. That the present results are probably stable is reflected in the results with a new sample of 69 male and female *Ss* in which it was found that TA and SD correlated to the extent of −.41. Further research dealing with this phenomenon appears indicated.

The negative correlation of TA with many intellectual measures leaves open the question of whether (*a*) high TA *Ss* get low scores on intellectual tests due to heightened anxiety while taking tests and not to intellectual deficiencies, or (*b*) persons who obtain high TA scores are simply lower in intellectual endowment than other *Ss* and therefore are more anxious about tests. Previous studies (Sarason, 1957), using the graphic rating test anxiety scale and Yale students as *Ss*, suggest that the first possibility may be closer to the actual state of affairs than the second. It was found that the negative correlation between test

anxiety and grade-point average approached a zero correlation as a function of number of years in college. This observation may imply that given sufficient time to adapt to perceived threatening situations, anxiety reactions may extinguish and that when this occurs, the correlation between test anxiety and intellectual measures becomes nonsignificant.

Although the results seem useful, most of the significant correlations obtained were quite low and their predictive power is minimal. Clearly, intellectual factors contribute most to intellectual performance. This certainly is not surprising, and probably the only conceivable practical role of tests like TA in large scale grade-point average prediction programs would be that of a suppressor variable.

The absence of significant relationships between general anxiety and intellectual measures is consistent with evidence already reported with the MAS (Sarason, 1956; Schulz & Calvin, 1955). It seems reasonable to conclude that MAS is unrelated to intellectual performance for undergraduate college students. However, it is not justifiable to conclude that this lack of relationship will hold for samples of *Ss* differing in various characteristics from those found in the present sample.

SUMMARY

Three hundred and nine men and 67 women were administered the following personality tests: (*a*) true-false test anxiety scale (TA), (*b*) Bendig's 20-item version of the Taylor Manifest Anxiety scale (MAS), (*c*) the Lack of Protection scale (LP), and (*d*) Edwards' Social Desirability Scale (SDS). These four variables and 13 measures of intellectual performance were intercorrelated.

TA correlations with MAS and LP for men and women ranged between +.36 and +.49. TA correlated −.23 for men and −.49 for women with SD. Both MAS and LP correlated much more negatively with SD than did TA. There was a tendency for all three anxiety measures to correlate more negatively with SD for women than for men.

Although all correlations between anxiety and intellectual measures were small, there was a consistent, significant tendency for TA to correlate negatively with the intellectual measures. No similar tendency was observed for the MAS, LP, and SDS.

REFERENCES

BENDIG, A. W. The development of a short form of the Manifest Anxiety Scale. *J. consult. Psychol.*, 1956, **20,** 384.

EDWARDS, A. L. *The social desirability variable in personality assessment and research.* New York: Dryden, 1957.

GORDON, E. M., & SARASON, S. B. The relationship between "test anxiety" and "other anxieties." *J. Pers.*, 1955, **23,** 317–323.

MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47,** 166–173.

SARASON, I. G. The relationship of anxiety and "lack of defensiveness" to intellectual performance. *J. consult. Psychol.*, 1956, **20,** 220–222.

SARASON, I. G. Test anxiety, general anxiety, and intellectual performance. *J. consult. Psychol.*, 1957, **21,** 485–490.

SARASON, I. G. Interrelationships among individual difference variables, behavior in psychotherapy, and verbal conditioning. *J. abnorm. soc. Psychol.*, 1958, **56,** 339–344.

SCHULZ, R. E., & CALVIN, A. D. A failure to replicate the finding of a negative correlation between manifest anxiety and ACE scores. *J. consult. Psychol.*, 1955, **19,** 223–224.

Received May 12, 1958.

UNCONSCIOUS SELF-EVALUATION USING A STEREOSCOPE[1, 2]

HALLA BELOFF AND JOHN BELOFF

Queen's University, Belfast

GALTON (1907, pp. 226–7) himself was aware of the fact that a stereoscope can be used to produce fusion between two *different* faces so that the observer sees only a single face, having some of the characteristics of each. However, Engel (1955a, 1955b, 1956) was the first to explore this situation in a controlled manner. He had established two facts relevant to this paper. First, if the eyes are presented with two different but compatible stimuli, as in the case of two faces, the subject will perceive a true synthesis and, so long as the faces are suitably matched for size and provided that he takes care to keep both eyes open

[1] Based on a paper read to the British Psychological Society, London, December, 1957.

[2] This research was conducted while Halla Beloff held a personal grant from the Medical Research Council, London.

all the time, he will suspect nothing amiss. Secondly, both Engel (1955a) and Drever (1957) report confirmation of the Galton finding that the composite face tends to be more satisfactory aesthetically than either of the component faces.

Many applications of this technique were suggested as early as 1877 (Galton, 1907) and still await investigation. However, one of the most interesting, perhaps, is the fusion of self with stranger. This composite we used to throw further light on the so-called problem of unconscious self-judgments, following the work initiated by Wolff (1943). We consider our technique to present another "personal product," analogous to mirror-reversed handwriting, for example, but more potent in the sense that the face is probably the center of our strongest self-cathexis. It will be remembered that Wolff found that when no conscious recognition had taken place, most subjects overvalued their own products, some undervalued them, and very few were neutral towards them. These findings have since been substantially corroborated by Huntley (1938, 1940) and Epstein (1955).

The hypothesis put forward was that if the subject does not recognize himself he will still give a differential response to a composite image incorporating his own face, relative to a stranger composite. Using ratings on the dimension of "attractiveness," we predicted that ratings would be differentially more favorable in the majority of cases.

In brief, our experimental design called for the rating of a control face and a rating of a self-stranger composite, matched for extraneous variables. There were two of these we considered especially important: 1. The fact that the subject may be more or less attractive than the comparison. In the selection of the control, therefore, *subjectively* equivalent faces were chosen from data obtained at a preliminary rating session. 2. The fact that, as we have mentioned, any composite is judged as more pleasing than its components considered separately. Thus, as the experimental face must consist of a composite, the control too was a fusion of two stranger faces.

METHOD

Subjects

Twenty-three men and 29 women attending a single elementary course in psychology at Queen's University served as *Ss*. The experiment took place during their first term; they may therefore be considered psychologically naive.

Material

A simple stereoscope mounted on a wooden screen (24″ x 30″) was used. Passport photographs of the *Ss* were obtained, unknown to them, from the University Registration Office.[3] These were reproduced to a standard specification, i.e., that the distance between the line joining the pupils and the line between the lips be held constant at 15 mm.[4]

Procedure

Preliminary rating session. The *Ss* were shown the photographs of 10 young strangers of their own sex (openly) and asked to rate them on a 7-point scale of attractiveness from +3 through −3. Attractiveness was defined as the joint product of good looks and a pleasing personality, insofar as the latter could be judged from a photograph. It could have been called a dimension of likeableness. The peculiarity of our rating scale was that *S* was to take himself as the standard, so that a positive rating meant that the stranger was considered more attractive than *S* by *S*, a negative rating that he was thought less attractive, while zero implied subjective equivalence. Three control faces were selected for each *S* from among those strangers who received a zero rating from him, or at worst fell into the plus or minus one category.

The experiment proper. This took place some four weeks after the initial rating session. No reference was made to the latter. Each *S* was first given a visual efficiency test using the Stereosette test and a simple monoptoscope.[5] The Stereosette test served mainly to give binocular fusion practice. The monoptoscope gave crude information about relative eye dominance. In the main task, *S* was presented with three double faces in the stereoscope. Card I contained the photographs of two randomly chosen strangers (same for all *Ss*), used purely for warming-up. Card II consisted of two control strangers, and Card III of self with the third control stranger. (The self photograph was always placed in the field of the nondominant eye.)

Instructions were as follows:

We are going to show you some photographs of faces in this apparatus and we want you to tell us what impression they make on you. Keep both eyes open all the time. This is absolutely essential. Tell us anything

[3] We acknowledge with thanks the cooperation of this office.

[4] In this way the general area of the faces was to some extent equated and the most important features, the eyes and mouth, in a position to be perceptually superimposed. "Fuzziness" that may appear in the rest of the face seems to be ignored by the *Ss*.

[5] Copyright for the latter is held by Stereosette Ltd., 48 Gresham Street, London E.C.2. A monoptoscope is a tube accommodating both eyes but converging to a point such that fixation of an object can be with one eye only. We are indebted to Robin Farr, a graduate student, for carrying out this part of the project.

TABLE 1
DISTRIBUTION OF RATINGS ON ATTRACTIVENESS

	−3	−2	−1	0	+1	+2	+3[a]
Control composite	0	1	3	13	12	5	0
Self-composite	0	0	2	8	11	12	1

Note.—Median test $\chi^2 = 3.76$ (.05 level = 3.84).
[a] −3 extremely unattractive, + 3 extremely attractive.

you can about the person. Finally rate each one on a scale of attractiveness, on a 7-point scale. But zero this time is to stand for complete indifference, for neutrality, +3 for high attractiveness with regard to the population in general, and −3 for extreme unattractiveness.

RESULTS

We found no case at all in the two initial cards of neutral photographs in which fusion did not spontaneously occur, or in which suspicion was aroused. In the third, crucial trial, fusion did sometimes break down and recognition occurred in about one third of the cases.[6] This left us with 15 men and 19 women.

Considering the 34 *S*s who did not consciously recognize themselves, their rating distributions for the control and self composite are shown in Table 1. Inspection of the table shows that a difference does exist between the two distributions, the mode of the control being at "0" and the mode of the self composite at "+2." The direction of this difference is in accord with our hypothesis.

Although these are not independent distributions, the Median test was applied. We consider that the lack of independence could work only against the hypothesis, any response bias making for greater similarity between the distributions.

As is seen, the chi square value very nearly approximates that necessary for rejection of the null hypothesis at the .05 level. Thus it would appear that the self-image is indeed preferred even above subjectively equivalent strangers when no overt recognition has occurred.

DISCUSSION

The exact dynamic background of nonrecognition in this type of situation is still a matter of controversy, and the present study was not designed to adjudicate between rival theories. Wolff (1943) postulated active psychic resistance to recognition and considered that the unconscious, in the Freudian sense, was responsible for the judgments. We ourselves favor the more parsimonious hypothesis of Huntley (1938), that failure of conscious recognition is due only to insufficiency of cues, although the cues are still adequate for the invocation of a differential self-response.

Wolff (1943) has stressed that unconscious self-judgments even if mainly favorable tend towards both extremes and avoid the neutral category. Our data like those of Epstein (1955), show no evidence of this. Instead, we have found a gradual increase from unfavorable to favorable ratings. Both with regard to the numerical ratings and the qualitative descriptions we found no dramatic differences between the control and critical stimuli. *S*s did not give a longer or deeper description of the self-composite, nor did they exhibit emotion when confronted with it, points about which Wolff has been very positive.

Our results though indicative are slender. One possible objection is that a self-stranger composite is not a simple disguised self-product at all, but at best looks like some plausible relative of the *S*. And after all, there is no law in psychology which says that we overesteem our relatives! Somewhat crudely, our answer is that somewhere in the visual cortex a complete projection of the self-photograph must exist, however overlaid. We investigated what the *S* made of this information.

The authors consider rather that the relatively small difference between the distributions is due to their especially stringent controls. Previous studies have always used essentially objectively equivalent controls, such as might have been obtained by having the self-products prerated by independent judges. However, we know that objective, external estimates of aspects of the self, whether concerned with personality or appearance, are often widely divergent from subjective judgments. Hence, we cannot know to what extent the overvaluation found under conditions of nonrecognition would not have been found at the conscious level also. At the peril of partialling out the very thing we were looking for, we decided to make the *S* himself responsible for the controls with which he would be matched. In summary, we would say that there exists not only a self-enhancement effect in conscious estimates but relative to them another enhancement in unconscious judgment. Hence, by separating the two we have a seemingly reduced effect.

SUMMARY

Each subject was presented with his own photograph in a stereoscope, together with that of a stranger, so that a single composite face was seen. Thirty-four subjects who did not recognize themselves under these conditions were then asked to

[6] In most cases, recognition was demonstrated by an immediate laugh and some exclamation like "That's me!" With a small minority, recognition was only established at the final enquiry, when *S* would admit that he had had a suspicion that he was somehow seeing himself, these *S*s, too, were of course discarded.

rate this self-composite for degree of attractiveness, and each subject's rating was compared with one that he made on a control composite consisting of two strange faces. Confirmation was obtained for the hypothesis that the self-composite would be rated higher than the control, thus lending support for the theory of the favorableness of unconscious self-judgments.

REFERENCES

DREVER, J. A binocular phenomenon in social perception. Paper read at British Psychologial Society, St. Andrews, April, 1957.

ENGEL, E. Binocular fusion of dissimilar figures. Unpublished paper, Princeton Univer., 1955. (a)

ENGEL, E. Meaningful content in the study of rivalry and fusion—preliminary observations. Unpublished technical report, Princeton Univer., 1955. (b)

ENGEL, E. The role of content in binocular resolution. *Amer. J. Psychol.*, 1956, **69**, 87–91.

EPSTEIN, S. Unconscious self-evaluation in a normal and a schizophrenic group. *J. abnorm. soc. Psychol.*, 1955, **50**, 65–70.

GALTON, F. *Enquiries into human faculty and its development.* London: Dent, 1907.

HUNTLEY, C. W. Judgments of self based upon records of expressive behavior. Unpublished doctoral dissertation, Harvard Univer., 1938.

HUNTLEY, C. W. Judgments of self based upon records of expressive behavior. *J. abnorm. soc. Psychol.*, 1940, **35**, 398–427.

WOLFF, W. *The expression of personality.* New York: Harper, 1943.

Received May 12, 1958.

THE EFFECT OF MOTIVATION AROUSAL ON HUMOR PREFERENCES[1]

JOHN F. STRICKLAND[2]

Johns Hopkins University

THE influence of different types of motivation on the appreciation of humor is investigated in the present study, which stems from the classic psychoanalytic formulations put forth by Freud (1938). In brief, he postulated that the pleasure derived from responding positively to humorous material is a result of gratification of unconscious wishes or motivations that would remain repressed under ordinary circumstances. In this sense, humor serves a real and important function of tension or drive reduction. The impulse that was formerly unacceptable, for example, aggression towards some authority figure, now becomes freely expressible as a pleasurable response to humorous material. The drive has now been channeled into an outlet that is now not only socially acceptable, but also acceptable to the individual himself without the concomitant anxiety that might otherwise attend the realization of these impulses.

In a recent psychoanalytically oriented approach to the study of humor, Redlich, Levine, and Sohler (1951) developed the Mirth Response Test, which has subsequently been modified by Doris and Fierman (1956). The test consists of a series of cartoons that have been rated by judges along the dimensions of aggression, sex, and nonsense. From a subject's ratings of the humorousness of each cartoon, his score on these three components can be obtained. A similar measure was used in the present study.

Although the Freudian approach to humor stresses repression almost exclusively, it is possible that the mechanisms of suppression and displacement are equally important or more important in the psychology of humor. The present experiment is intended to provide evidence on this point.

If the subject's responses are sensitive enough to be influenced by momentary suppressions and displacements, then a careful manipulation of environmental factors should produce definite humor preferences in predictable ways. It is specifically hypothesized that a person who is annoyed and frustrated but is prevented from expressing his hostility should respond more favorably to humorous material of a hostile nature. Similarly, a subject who is aroused sexually but has no means of expressing this motivation should prefer humor material with sexual content.

METHOD

The technique of measuring humor preferences used in this experiment is very similar to that used by Doris and Fierman (1956). Thirty-three cartoons were selected so that 11 represented hostile

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements for the degree of doctor of philosophy in the Department of Psychology of Johns Hopkins University. The author wishes to express his deepest gratitude to James Deese who acted as faculty advisor to this dissertation and to Clinton DeSoto who served as referee. Their assistance and suggestions were invaluable.

[2] Now at the U. S. Navy Electronics Laboratory, Human Factors Division, San Diego 52, California.

or aggressive themes, 11 represented sexual themes, and 11 represented neutral or nonsensical themes. The cartoons were selected from a collection of over 150. Fifteen judges, including faculty members, graduate students, and undergraduates, sorted these samples into four categories according to the general theme of the cartoon: sex, aggression, nonsense, and a mixed category for cartoons that could not definitely be placed in one of the other three. A cartoon was included in the final selection only if all 15 judges had unanimously placed the cartoon in the same category. In the experiment, the cartoons were presented to each subject (*S*) individually in a different random order with instructions that he was participating in the standardization of a humor test designed to find out how funny these cartoons were. The *S* was asked to rate each cartoon on a 15-point graphic scale that ranged from not funny to extremely funny. From these ratings a mean score was obtained for each *S* for each of the three types of cartoons—hostile, sexual, and neutral.

Subjects

Seventy-five students enrolled in the introductory psychology course at Johns Hopkins University served as *S*s, as part of their regular course requirements. Twenty-five *S*s were assigned at random to each of three experimental groups. Each *S* was tested individually, and a measure of his humor preference recorded.

Conditions

A mixed analysis of variance design was employed in which all *S*s in each of three experimental conditions (chosen to arouse different types of motivation) were given the same three-component humor test.

In the first experimental condition, the experimenter (*E*) attempted to arouse hostility and anger in the *S*s (Hostile group). When the *S* appeared at the appointed time, the *E* conveyed through his tone of voice and small gestures his displeasure at being interrupted. The *S* was then ushered curtly into an adjoining room and told that the *E* would be with him "in just a minute." The room itself was windowless and completely bare except for a straight chair and table. The door was then slammed shut and the *S* was left alone for nearly 20 minutes. At the end of this period the *E* re-entered the room and without apology or explanation administered the humor test.

The second condition (Sexual group) attempted to arouse sexual motivation following an approach suggested by Clark (McClelland, 1955). When the *S* was introduced into the experimental situation he was presented with 10 8 x 10 photographs of nude photographer's models. He was told that the pictures had been rated by art students on their artistic qualities and that we were trying to determine the relationship between art and sex; it was then his job to rate the pictures solely on their sexual attractiveness. He was asked to rank the 10 photographs in order of his preference and then in a paragraph to describe what he particularly liked about his first choice. When this task was completed, the *S* was given the humor test.

In the third condition (Control group) the *S* was given the humor test immediately after being brought into the experimental situation.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF MEAN HUMOR RATINGS OF CARTOONS

Source of Variation	Sum of Squares	*df*	Mean Square	*F*
Between Subjects	718	74		
Conditions (3 groups)	36	2	18	1.89
Error (Between *S*s wG)	682	72	9.5	
Within Subjects	261	150		
Cartoons (Hostile, Sexual, Neutral)	5	2	2.50	1.57
Cartoon × Condition Interaction	26	4	6.50	4.09*
Error (Cartoon × *S*s wG)	230	144	1.59	
Total	979	224		

* $p < .005$.

RESULTS

Table 1 summarizes the results of the analysis of variance performed on the data. The scores used in the analysis were the mean ratings, for each *S*, of each of the three types of cartoons.

Table 2 gives the means and *SD*s of the three types of cartoons for the three conditions, and also indicates which pairs of means were significantly different. The Critical Difference (Lindquist, 1953, Chap. 3) computed at the .01 level was found to be .31. Here the internal differences reflected in the significant interaction are quite clear. Within the Hostile group, the hostile cartoons were clearly preferred to both sexual and neutral types ($p < .01$) and within the Sexual group the sexual cartoons are clearly preferred to both hostile and neutral ones ($p < .01$). This latter finding may be tempered somewhat by the fact that within the Control group the sexual cartoons were preferred significantly more than the other two types of cartoons. One other finding to be discussed is that neutral cartoons in the Sexual group were more strongly preferred than the sexual and the neutral cartoons in the Hostile group.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF HUMOR SCORES FOR EACH CONDITION

	Hostile Group			Sexual Group			Control Group		
	H	S	N	H	S	N	H	S	N
Mean[a]	7.79	6.61	6.73	7.73	8.44	7.78	7.17	7.52	7.13
SD	3.30	3.14	2.65	3.05	3.40	2.89	3.19	3.40	3.07
Grand Mean		(7.04)			(7.98)			(7.27)	

Note.—Significantly Different Means:[b]
HH → HS, HN, CH, CN
SH → HS, HN, CH, CN
SS → HH, HS, HN, SH, SN, CH, CS, CN
SN → HS, HN, CH, CN
CH → HS, HN
CS → HS, HN
CN → HS, HN

[a] Critical Difference at .01 level = .31.
[b] The first letter refers to the experimental condition and the second letter refers to the type of cartoon within the condition. The mean to the left of the arrow is significantly greater than all those to the right of the arrow.

DISCUSSION

The original hypothesis that *S*s placed in hostile and sexually arousing situations should prefer hostile and sexual cartoons, respectively, is borne out by the significant interaction. This finding is important in two ways. First, it demonstrates how sensitive the appreciation of humor is to situational influence and how, by separately arousing different types of motivation, the *S*'s response to humorous material can be controlled to a considerable extent. From Table 2 it can be seen that the effect of the hostile condition was to decrease the responses to the sexual and neutral cartoons compared to the Control group, while the effect of the sexually arousing situation was to elevate the responses to the sexual and neutral cartoons. The response to the hostile cartoons for these two conditions was approximately the same. Thus it would appear that the hostile situation was much more antagonistic to the appreciation of humor, with the exception of hostile humor, while the sexually arousing situation allowed for the appreciation of other humorous material at the same time sensitizing the person toward a greater appreciation of material with sexual content. That is, the sexual motivation did not conflict with the appreciation of humor.

The second point concerns the Freudian approach to the theory of humor. It can be argued that if repression were truly successful as a defense, then it would probably operate even with humorous stimuli. For example, if the *S* were truly repressing hostility and aggression he would very likely respond negatively to aggression in humor or actually to "miss the point" of aggressive jokes. If, on the other hand, aggression is simply suppressed or restrained by the nature of the situation, suppression may be the more important factor. While the present experiment cannot rule out the possible operation of repression, it tends to support a suppression hypothesis, since the results can be adequately explained without calling on repression. Both experimental conditions can be viewed as arousing motivations in the *S* that he must simply suppress or restrain. In the hostile condition, the *S*s were usually quite visibly angered, yet the *E* was enough of an authority figure to prevent any overt expression of the anger; in the sexually arousing situation, the task was carried out as a serious experiment which the *S*s could not make light of, yet provided no outlet for sexual motivation. Thus the sexual and aggressive motivations that were suppressed in the experimental conditions were expressed in increased appreciation for the sexual and aggressive content in the humor material.

While the present results may be adequately explained more economically in terms of suppression rather than of repression, repressed motivations might account for an individual's consistent trend of humor preferences that operated in so-called "neutral" situations. Suppression, on the other hand, may account for momentary short-term preferences that can be superimposed when specifically provoked. Thus a person with a general inclination to favor jokes with a sexual theme might prefer hostile wit if placed in a sufficiently frustrating situation. Periodic testing over long-range intervals might provide some evidence on this point.

SUMMARY

The present experiment was designed to test the hypothesis that the arousal of different types

of motivations would produce differences in humor preferences. It was predicted that in a hostility-arousing situation subjects would prefer humorous material of a hostile and aggressive nature, while subjects placed in a sexually arousing situation would prefer humorous material with a sexual theme. This prediction was confirmed by the experiment. An explanation of the results was offered in terms of suppression.

REFERENCES

DORIS, J., & FIERMAN, E. Humor and anxiety. *J. abnorm. soc Psychol.*, 1956, **53**, 59–62.

FREUD, S. *Wit and its relation to the unconscious. The basic writings of Sigmund Freud.* New York: Modern Library, 1938.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

MCCELLAND, D. C. (Ed.) *Studies in motivation.* New York: Appleton-Century-Crofts, 1955.

REDLICH, F. C., LEVINE, J., & SOHLER, T. P. Mirth Response Test: Preliminary report on a psychodiagnostic technique utilizing dynamics of humor. *Amer. J. Orthopsychiat.*, 1951, **21**, 717–734.

STRICKLAND, J. F. The effect of motivation arousal on humor preferences. Unpublished doctoral dissertation, Johns Hopkins Univer., 1958.

Received May 15, 1958.

THE EFFECTS OF HYPNOTICALLY INDUCED ANXIETY ON THE MANIFEST ANXIETY SCALE AND THE BARRON EGO-STRENGTH SCALE

HANUS J. GROSZ

Institute of Psychiatric Research, Indiana University Medical Center[1]

AND EUGENE E. LEVITT

Indiana University Medical Center

A WEAKNESS of the collected work bearing on the validity of Taylor's Manifest Anxiety Scale (MAS) as a measure of clinical anxiety is that it consists almost entirely of correlational studies. These studies have usually yielded positive results, and there seem to be enough of them so that we may reasonably conclude that the MAS has good construct validity. Nonetheless, there appears to be considerable skepticism concerning the test's general validity, as witness the statement of Taylor herself (1956) and the criticisms of Jessor and Hammond (1957). This conclusion applies also to the Barron Ego-Strength Scale (Es), which has so far been successful in differentiating psychiatric groups from normals (Barron, 1953; Taft, 1957; Tamkin & Klett, 1957).

Studies in which the variable has been manipulated in order to determine the effects on test performance are less likely to be subject to this type of confounding, and are therefore usually more convincing tests of validity, depending largely on the soundness of the definition of the factor as an independent variable. Gallagher (1953) administered the MAS to patients before and after undergoing psychotherapy, and Lebo, Toal, and Brick (1958) tested before and after CO_2 treatment. In both studies, significant decreases in MAS scores were reported. Barron and Leary (1955) report no mean difference between treated and untreated cases on the Es, though Cartwright (1956) points out that there was a significant difference in the variability of scores for the two groups.

In the present study, anxiety, independently defined, was induced in normal subjects, and pre- and postinduction MAS and Es scores were obtained. Anxiety was induced by means of hypnotic suggestion. A number of studies (Counts & Mensh, 1950; Eichhorn & Trachtir, 1955; Levine, Grassi, & Gerson, 1943; Sarbin, 1939) have indicated that this method can be very successful in inducing various emotional states. It seemed highly effective in this investigation.

METHOD

The subjects (*Ss*) were 12 medical and nursing students, 8 females and 4 males. The age range was 18–29 years with a mean of 21.5. All were paid volunteers, selected from a pool of volunteer *Ss* on the basis of hypnotizability and psychological adjustment. None had any serious personality difficulties or problems as determined by a psychiatric examination.

Since the state of hypnosis itself could have some effect on test scores, the MAS and the Es were administered to each *S* three times: in the ordinary waking state, in a hypnotic trance, and in hypnotically induced, diffuse anxiety. Four *Ss*

[1] Now in the Department of Medicine (Neurology), Albert Einstein College of Medicine.

completed the tests in that order. The anxiety state administration preceded the hypnotic state administration for three *S*s; the order of administration for the remaining five *S*s was hypnosis, anxiety, waking state. The variations in order of administration were introduced to neutralize any possible effects of order on test scores.[2] *S*s were not made amnesic for previous administrations because it did not seem necessary.

The suggestions for inducing anxiety were an elaboration of the type used by Levine, Grassi, and Gerson (1943):

I suggest to you that you will begin to feel in a certain way, that you will have certain definite feelings which I am going to suggest to you. You will begin to feel these feelings and emotions so vividly that your whole person, your whole body, every fiber of yourself, will feel that way—at first slightly, then gradually stronger and stronger. These feelings will last until I suggest to you that they are no longer there. The feelings and emotions which you will begin to experience are those of anxiety—feelings of anxiety, more and more intensely, more and more vividly. You will become more and more anxious. At first, there will just be a feeling of apprehension. Then, gradually you are beginning to feel more and more afraid but you do not know what it is that makes you feel so afraid and so anxious. You just feel that way without knowing why. You feel more and more afraid and anxious all the time, as if something dreadful is going to happen to you. Yet, you do not know what this dreadful thing is. I suggest that you will gradually begin to experience a state in which you feel really panicky, where you are so fearful that you feel an almost unbearable dread, fear, and panic. In fact, you are most likely already feeling this way and you will continue to feel this way more and more strongly.

The suggestions were delivered with increasing vocal intensity and emphasis. The use of a number of general synonyms for anxiety—fear, apprehension, etc.—represents an effort to circumvent the effects of different private meanings of the single word "anxiety" among *S*s.

After induction, all *S*s verbalized anxiety feelings, and all showed some physical manifestations of anxiety like trembling, agitation, pallor, etc. Ratings of degree of anxiety were made independently by a psychiatrist and a psychologist (the authors) for each case. A 5-point scale was used, a rating of 1 representing no anxiety (as in the waking state) and those from 2 to 5 representing increasing degrees of anxiety. The interrater agreement, as reflected in a tau coefficient of .47, is not high, but the raters did agree that every *S* showed *some* anxiety. The mean rating by the psychiatrist was 3.04, $SD = .82$. The psychologist's mean rating was 2.92 with an SD of .83.

[2] The method of inducing anxiety was somewhat different for these last five *S*s in that the *E* repeated various sections of the anxiety-inducing suggestions during test performance. This precludes the computation of a variance due to order of administration in the variance analyses of the next section, since order and technique are confounded. The MAS scores of these *S*s tended to be higher in the anxiety state than those of the remaining *S*s.

TABLE 1
MEANS AND *SD*s ON THE MAS AND Es IN THE THREE STATES

State	MAS		Es	
	Mean	*SD*	Mean	*SD*
Waking	12.08	3.99	46.17	3.56
Hypnosis	12.67	6.01	43.75	4.01
Anxiety	25.00	10.77	39.33	5.46

$F_{MAS} = 12.08$, $p = <.001$.
$F_{Es} = 10.34$, $p = <.001$.

RESULTS

The results are summarized in Table 1.

The scores for the three different states for both tests were subjected to a variance analysis in which the main factors are states and subjects. The apparent heterogeneity of the variances of individual state distributions, as reflected in the *SD*s of Table 1, is difficult to assess because the distribution means are correlated. It would not, however, as Lindquist (1953) points out, preclude the variance analyses. As a precaution, however, significance at the .025 level, as recommended by Lindquist, was required.

For both scales, the variances due to states were significant beyond the .001 level. The t tests between pairs of state means show that this significance in both instances was due to differences between the anxiety state and the other two means. The means for the waking and hypnotic states did not differ. For the MAS, the t values were 0.20 for waking vs. hypnosis, 4.35 for waking vs. anxiety, and 4.15 for hypnosis vs. anxiety. For the Es, the respective t's were 1.59, 4.49, and 2.90.

In an attempt to estimate the consistency of the data from state to state, the scores were ranked and Kendall's coefficient of concordance, W, was computed for each scale. For the MAS, W, $= .37$, $p = .30$. For the Es, $W = .64$, $p = < .05 > .02$.

Rank order correlation coefficients between the two scales were .08 for the waking state, .02 for the hypnotic state, and $-.50$ for anxiety. None of these values is significant. The absence of any relationship may reflect the limited range of

scores which results from a sample screened for abnormality.

DISCUSSION

The important finding seems to be that hypnotically induced anxiety resulted in a significant increase in MAS scores and a significant decrease in Es scores. Only two *S*s failed to show increases or decreases in the right direction, and one *S* showed no change on the Es.

The *W*s may not be regarded as reliable estimates of concordance among the three states because of the difference in the technique for inducing anxiety among the *S*s (see Footnote 2). This discrepancy would tend to produce greater changes in scale scores in anxiety relative to waking and hypnosis state scores for some *S*s and would, therefore, reduce the rank correlations. However, since this discrepancy is a constant for both scales, a comparison between *W*s is not likely to be biased. It should be noted that the *W* for the MAS was insignificant, whereas that for the Es is significant beyond the .05 level. This suggests that the Es is a more stable measure than the MAS, which is just what would be expected of the scales if they are really measuring ego-strength and anxiety. This comparison of *W*s is supported by the observation that the MAS mean in anxiety is over 2 *SD*s higher than in hypnosis, and the Es mean in anxiety is slightly more than 1 *SD* lower than in hypnosis. In other words, anxiety affected performance on the MAS to a considerably greater extent than it did the Es.

These results suggest that the MAS is a valid indicator of clinical anxiety. They also suggest that the Es is genuinely measuring ego-strength or some similar trait associated with psychopathological conditions. These conclusions must be tempered by the small, select sample used, but when added to earlier findings, especially with the MAS, they seem quite convincing.

REFERENCES

BARRON, F. An ego-strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, **17**, 326–333.

BARRON, F., & LEARY, T. F. Changes in psychoneurotic patients with and without psychotherapy. *J. consult. Psychol.*, 1955, **19**, 239–245.

CARTWRIGHT, D. S. Note on "changes in psychoneurotic patients with and without psychotherapy." *J. consult.Psychol.*, 1956, **20**, 403–404.

COUNTS, R. N., & MENSH, I. N. Personality characteristics in hypnotically induced hostility. *J. clin. Psychol.*, 1950, **6**, 325–330.

EICHHORN, R., & TRACHTIR, J. The relationship between anxiety, hypnotically induced emotions and gastric secretion. *Gastroenterology*, 1955, **29**, 422–431.

GALLAGHER, J. J. Manifest anxiety changes concomitant with client-centered therapy. *J. consult. Psychol*, 1953, **17**, 443–446.

JESSOR, R., & HAMMOND, K. R. Construct validity and the Taylor Anxiety Scale. *Psychol. Bull.*, 1957, **54**, 161–170.

LEBO, D., TOAL, R. A., & BRICK, H. Manifest anxiety changes in prisoners before and after CO_2. *J. consult. Psychol.*, 1958, **22**, 51–55.

LEVINE, KATE N., GRASSI, J. R., & GERSON, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach: A case study. *Ror. Res. Exch.*, 1943, **7**, 130–144.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. New York: Houghton Mifflin, 1953.

SARBIN, T. R. Rorschach patterns under hypnosis. *Amer. J. Orthopsychiat.*, 1939, **9**, 315–321.

TAFT, R. The validity of the Barron Ego-Strength Scale and the Welsh Anxiety Index. *J. consult. Psychol.*, 1957, **21**, 247–249.

TAMKIN, A. S., & KLETT, C. J. Barron's Ego-Strength Scale: A replication of an evaluation of its construct validity. *J. consult. Psychol.*, 1957, **21**, 412.

TAYLOR, JANET A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 304–320.

Received June 6, 1958.

PRIMARY PROCESS IN CHILDREN AS A FUNCTION OF AGE[1, 2]

ALVIN G. BURSTEIN

University of Chicago

AN AREA of concern in developmental psychology is the problem of characterizing the difference between the thinking of younger and older children. Freud (1956) makes such a developmental distinction in terms of his concept of primary and secondary process when he states:

> The primary processes are present in the mental apparatus from the first, while it is only during the course of life that the secondary processes unfold, and come to inhibit and overlay the primary ones; it may even be that their complete domination is not attained until the prime of life (p. 603).

Freud (1949) further specified what would appear to be an experimentally accessible aspect of primary process thought saying "... contraries are not kept apart from each other, but are treated as though they were identical" (p. 53). In addition, Freud relates this confusion of opposites and identicals specifically to the area of language. In a discussion of the role of opposites in dream interpretation, he remarks: "... the most ancient languages behave exactly like dreams in this respect. In the first instance, they have only a single word to describe two contraries ... (strong-weak, old-young, far-near, bind-sever) ..." (p. 318n).

These remarks by Freud suggest that a measure of the tendency to confuse opposites and identicals might be the degree of difficulty experienced in distinguishing antonyms from synonyms. Since it results from primary process activity, this difficulty would be greater in the case of younger children.

Such a prediction might also be made upon other than psychoanalytic bases. For example, rather than assume that different modes of mental function exist (i.e., primary and secondary processes), one might speculate that whenever concepts are learned it is by the gradual differentiation of polar opposites out of an undifferentiated primordial concept, a process reminiscent of that suggested by Werner (1948). That is, one might think of progressive stages of increasingly precise knowledge concerning a concept or word and, consequently, predict differing degrees of difficulty in distinguishing opposites as a result of differing degrees of "education" on the part of children of differing ages.

[1] This paper is based upon a thesis submitted in partial fulfillment of the requirements for the degree of master of arts, University of Chicago, 1958. The author is indebted to his sponsor, Loren J. Chapman, for his valuable suggestions and criticisms at every stage of the research.

[2] The author is indebted to Joseph J. Engandela of Parkview Elementary School, Carpentersville, Illinois, for his cooperation in making subjects available.

PROCEDURE

A paper-and-pencil test of 97 items was constructed. On each item *S* was instructed to select the synonym for a given word from among three alternatives, consisting of a synonym, an antonym, and an irrelevant word. An example will serve to illustrate:

> Circle the word that means almost the same as BIG:
>
> small (ant.) individual (irr.) large (syn.)

The irrelevant word was chosen on the basis of its having a minimal relationship to the stimulus word. Its function was to measure choices made upon uncontrollable bases, e.g., obscure communalities, idiosyncratic associative connections, lack of knowledge, random marking, etc. Since such choices would by chance be equally divided among the three types of alternatives, the number of irrelevant markings is a measure of antonym markings which were made on such uncontrollable bases. Consequently, an error score of antonym markings minus irrelevant markings (A − I) is a relatively pure measure of the distracting effect of the antonyms on this task, that is, a measure of the degree to which synonyms and antonyms are confused.

As we have mentioned, choices made upon uncontrollable bases should be divided equally among the three types of alternatives. On the basis of an argument parallel to that of the above paragraph, a measure of the number of synonyms chosen on the basis of their being synonyms would be synonym markings minus irrelevant markings (S − I). Total errors would then be the difference between total choices (A + S + I) and synonym choices as such (S − I). Total errors ([A + S + I] − [S − I]) are numerically equal to antonym markings plus twice the irrelevant markings, i.e., (A + 2I). It is clear that "total errors" here does not mean total erroneous markings but, rather, the sum of total erroneous markings plus those synonym markings made for erroneous reasons.

In addition to the difference score (A − I), a

TABLE 1
COMPARISON OF ERROR SCORES FOR THIRD AND SIXTH GRADERS[a]

	Third Grade				Sixth Grade			
	A	I	A − I	$\frac{A - I}{A + 2I}$	A	I	A − I	$\frac{A - I}{A + 2I}$
Median	30.00	8.50	17.50	.40	8.25	3.38	2.25	.23
Mean	33.34	9.38	24.04	.41	12.41	6.07	6.33	.29
SD	18.04	4.38	8.76	.24	3.34	6.52	8.83	.38

[a] "A" refers to antonym choices; "I" refers to irrelevant choices. For an explanation of the compound scores, see text.

proportion score, such as the ratio of antonym choices as such to total errors (A − I/A + 2I) might be computed. This ratio gives the proportion of total errors attributable to the choice of an antonym on the basis of its being an antonym.

If, in the case of the younger children, all error tendencies are greater, and, in addition, no shift occurs in the strength of the tendency to confuse antonyms and synonyms relative to all other error tendencies, then one would expect the difference score to be higher in the case of the younger children, while the proportion score would be the same for both groups. If, on the other hand, there exists a differential increase (with respect to all other error tendencies) in the tendency to confuse antonyms and synonyms, then both the difference and the proportion scores should be greater in the case of the younger children, showing an increased tendency to be distracted by antonyms that cannot be accounted for on the basis of a general increase in all error tendencies.

The types of alternatives were roughly equated for frequency of occurrence in the language using Thorndike and Lorge's (1944) summary of four different word frequency counts.

The hypothesis may now be worded in operational terms, viz.: that both difference (A − I) and proportion (A − I/A + 2I) scores for younger children are greater than for older children.

The test was administered to 26 third-graders and 27 sixth-graders at Parkview Elementary School, Carpentersville, Illinois. The younger group consisted of equal numbers of children of both sexes, while the older group contained 14 males and 13 females.

RESULTS

The median, mean, and standard deviation for the error scores are presented in Table 1. The difference between the groups in terms of both difference and proportion scores was evaluated using a two-tailed Mann-Whitney *U* test. In the case of the difference score, the younger group received the higher error scores ($p < .0001$); the younger group also received the higher proportion scores ($p < .11$).

The difference scores made by the males were compared to those of the females by the same statistical test, and the difference was not significant ($p = .84$).

Since the largest number of irrelevants chosen by any *S* was 23, and the mean number of irrelevants chosen by both groups (9.38 and 6.07) differs significantly ($p < .0001$) from 32, the number expected by chance, one can minimally assume that the children did not mark the test forms randomly or in response to position of alternatives, and that they understood the mechanics involved in the task.

DISCUSSION

The clear experimental finding is that the younger children have a greater tendency to confuse antonyms with synonyms. In addition, there is some indication that this tendency cannot be attributed to a general increase in errors of all kinds.

These results can profitably be viewed in the light of research in the area of thought pathology by Chapman (1958) who, using a conceptual task, found that associative distracters based on contiguity of words, contiguity of objects, and rhymes, induced greater error tendencies among schizophrenics than among normal *S*s. That antonyms have associative "pull" is clear from the frequency with which a relationship of contrast exists between the stimulus word and the response in the Kent-Rosanoff (1910) word association lists. Hence, antonyms might constitute one case in the general class of associative distracters.

However, Woodrow and Lowell (1916), in comparing the responses of children and adults to the stimulus words of the Kent-Rosanoff list, found that adult responses classifiable as contrasts occur more often than the same type of response in the case of children, while responses classifiable as similarities occur with approximately equal frequency in both groups (adults, contrast—10.6%, similarities—8.9%; children, contrasts—1.3%,

similarities—8.6%). Consequently, although children may or may not be more susceptible to associative distracters than are adults, the findings of Woodrow et al. suggest that the pull of contrasts as a type of associative distracter relative to the pull of other types would be less in the case of children. Therefore, with respect to the present task, the associative relationship between the stimulus word and its antonym may be insufficient, of itself, to explain the greater tendency of the younger children to confuse antonyms and synonyms.

A possible line of research would be the application of an instrument similar to the present one in pathological cases. Such an application would appear especially appropriate in the case of schizophrenics, not only on the basis of psychoanalytic theory, but also in the light of clinical observation, including the report of Laffal, Lenkoski, and Ameen (1956) concerning the use of opposite speech by a schizophrenic patient.

REFERENCES

CHAPMAN, L. J. Intrusion of associative responses into schizophrenic conceptual performance. *J. abnorm. soc. Psychol.*, 1958, **56,** 374–379.

FREUD, S. *An outline of psychoanalysis.* New York: Norton, 1949.

FREUD, S. *The interpretation of dreams.* New York: Basic Books, 1956.

KENT, GRACE H., & ROSANOFF, A. J. A study of association in insanity. *Amer. J. Insanity,* 1910, **67,** Nos. 1 and 2.

LAFFAL, J., LENKOSKI, L. S., & AMEEN, L. "Opposite speech" in a schizophrenic patient. *J. abnorm. soc. Psychol.*, 1956, **52,** 409–413.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Columbia Univer. Press, 1944.

WERNER, H. *Comparative psychology of mental development.* New York: Follett, 1948.

WOODROW, H., & LOWELL, FRANCES. Children's association frequency tables. *Psychol. Monogr.*, 1916, **22,** No. 5 (Whole No. 97).

Received July 14, 1958.

ATTITUDES AND THE AVAILABILITY OF KNOWLEDGE CONCERNING TRADITIONAL BELIEFS

LEONARD W. DOOB

Yale University

Investigator: Why don't you cut your baby's hair?
Jamaican Woman: Because me fearful.
I: Why are you fearful?
JW: People say, when you do, him be dumb or talk late.
I: Do you think so, do you think that will happen?
JW: No.
I: Then why don't you cut it?
JW: Me fearful, me don't know, maybe them right.

BY DEFINITION, a traditional belief is not only known by most members of a society but also secures their approval. It is obviously possible, however, to be acquainted with a belief without assenting to it. Most Americans, for example, can name a number that is supposed to be unlucky and yet deny that 13 inevitably brings misfortune. For contact with a prevailing belief can scarcely be avoided; similar stereotypes, it has frequently been observed (e.g., Harding, Kutner, Proshansky, & Chein, 1954, pp. 1023–1024), pervade all sections of a population. In contrast, attitudes toward a belief vary with experience and consequently fluctuate, to a greater degree than knowledge, from group to group and from individual to individual within a society.

There must be, nevertheless, an association between knowledge and attitude concerning traditional beliefs, just as relations between the cognitive and affective aspects of attitudes have been demonstrated (Rosenberg, 1956). A favorable attitude signifies the functional utility of the belief to the individual; a useful belief is frequently evoked; hence, the threshold of such a belief is likely to be low. Being reminded of a traditional belief which has been rejected, moreover, is disquieting (Could 13 possibly bring bad luck?); by repressing and neglecting a rejected belief, consequently, dissonance between the apparent or alleged assent of one's peers or ancestors on the one hand and one's own dissent on the other can be reduced (Festinger, 1957). The causal sequence

may also be reversed. Frequent and perhaps involuntary contact with a traditional belief probably subjects the person to supporting arguments which often accompany it. Or that contact becomes disruptive unless one agrees with the implications of the belief and perhaps acts accordingly. At any rate, all these considerations lead to the hypothesis being tested: *traditional beliefs that evoke assent from a person are more readily available to him than those evoking dissent.* This report also explores other correlates of availability of knowledge and tests different methods of interviewing in another culture.

METHOD

The two studies reported here were based upon interviews obtained in Jamaica, B.W.I., during the summer of 1957.[1] Traditional beliefs are those identified by previous investigators (Carr, 1952; Taylor, 1955) or by competent Jamaican informants and then so labeled by a panel of Jamaicans. The *availability* of a belief was determined first by assuming that each informant had in fact been exposed to the belief in the past and by using what then can be called the method of recall. The informant was given a simple, open-ended question; for example, "What will happen if a woman with child drinks a lot of milk?" *Assent* was measured by asking him immediately whether or not he agreed with the belief which he himself had stated (if his reply had been "correct" from the standpoint of tradition) or which the investigator supplied (if his had been "incorrect" or not forthcoming.) An "incorrect" reply to the question in the illustration just given included any reference to the health of the pregnant woman, in which case the investigator said, without appearing to be offering a correction, "Some people think that, if a woman with child drinks a lot of milk, the child will be born light—what do you think?"

The first study is based upon 65 male and female adults who—in a sample of 112 interviewees—were tested for availability of knowledge. Except for 7 female social workers in Kingston, randomly selected subsamples came from the lower social and economic strata: 16 from a depressed rural area in the parish of Clarendon, 30 from a poverty-stricken housing development near Kingston, and 12 workers (full- or part-time fishermen) from a relatively prosperous area beside the sea in the parish of St. Ann. The data were collected individually during a two-hour interview which employed for other purposes a schedule of questions and tests adapted from ones previously administered to samples in three African societies (Doob, 1957). Ten traditional beliefs were tested.[2]

The second study, concentrating exclusively on the hypothesis which in fact occurred to the investigator while interviewing the adults, used 93 Jamaican children between the ages of 11 and 15 who were drawn from the highest forms of an elementary school in a semiurban area of St. Ann. They were interviewed, also individually, in an isolated room during school hours. Twenty traditional beliefs were tested.

Experimental variables were introduced into both studies. One involved the *context* in which both availability and assent were measured. The adults in the first study were randomly assigned to two groups. In one, the question was phrased as illustrated above (viz., "What will happen when . . .?") and, when the informant did not give the "correct" reply, the investigator supplied it without attribution. In the other, the phrasing was "What do *some people* say will happen when . . .?" and the "correct" reply, when it had to be supplied, was also attributed to "some people." In the second or school study, the first five questions were prefaced with "What do *some people* think will happen when . . .?"; each of the second five began with "What do *you* think will happen when . . .?"; and *you* was replaced by *intelligent, respectable people* and by *ignorant, superstitious people* in the third and fourth group of five questions respectively. The order of presenting the contextual variable to the children remained unchanged; instead, the order of the four sets of five beliefs was altered so that each child was randomly assigned to one of four experimental groups in a latin-square design. Once again, if the "correct" reply was not given by the child, it was provided by the investigator and ascribed to the appropriate "people"; for the second variable ("What do *you* think . . .?"), the ascription was to "some people."

Another variable was introduced into the second study, that of *inquiry*. Half the children selected on a random basis were asked to supply reasons for their agreement or disagreement with the

[1] The investigation was supported by a research grant (M-1940) from the National Institute of Mental Health, of the National Institutes of Health, U. S. Public Health Service. The writer is also deeply grateful to the following hospitable Jamaicans: Clifford E. Atkinson, Archdeacon Lennon, Pat Levy, Enid Jones Maier, Olive Sharp, and Kenneth Standard.

[2] Lists of beliefs, a table of phi coefficients, and three tables containing the data on which the p values mentioned in the Results section are based have been deposited with the American Documentation Institute. Order Document No. 5967 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

traditional belief but only in connection with the first set of five beliefs (introduced with the ascription to "some people") and the third set (introduced with the ascription to "intelligent, respectable people"). Reaction times to the knowledge and attitude questions were unobtrusively noted on a stop watch.

RESULTS

Results may be considered as a whole by disregarding the experimental variables of context and inquiry. For each belief a 2 × 2 table was constructed by dichotomizing availability of knowledge ("correct" answer vs. "incorrect" answer plus "don't know") as well as attitude ("agree" vs. "disagree" plus "don't know"). Then the degree of association was measured by means of phi coefficients. For all 10 beliefs among the adults and for all 20 among the children, that coefficient was positive; hence, a sign test suggests a clearly significant trend. One half of the 10 coefficients in the first study and also one half of the 20 in the second were themselves significant at the .05 level or less (one-tailed tests). Most of the significant coefficients were in the .30's; a few were higher. There was no discernible connection between the magnitude of the coefficients and the content of the beliefs, which—in the case of the children—had been deliberately selected to include different areas such as the moon and stars, child-rearing practices, and evil spirits.

With a second method of analysis the results can be approached from the viewpoint of the society as a whole: did those beliefs which ranked high with respect to assent also rank high with respect to availability? In both studies the percentages recalling "correctly" and agreeing were calculated for each belief; the beliefs were then ranked in both respects; and rho's were determined. For the adults, the figure was +.33 and is not significant; for the children it was +.75 and is significant.

When the same procedures were employed to appraise the subgroups established on the basis of the experimental variables of context for the adults and of context and inquiry for the children, the hypothesis was generally vindicated under the varying conditions. Identical trends with respect to positive phi's appeared; but the contextual variation of "What do *some people* think will happen when . . .?" among the children inexplicably produced a nonsignificant trend. Fewer of the phi's, moreover, were individually significant, no doubt in part because the *n*'s were correspondingly smaller. The rho's were also positive, but not all were significant.

The effects of the experimental variables can also be viewed in their own right. Among the adults of the first study, the variable of context had no effect whatsoever upon the availability of knowledge or the expression of attitude. When an analysis of variance in the second study was made by eliminating at random five cases in order to equalize the number in each cell, however, it was clear that both availability of knowledge and expression of attitude, besides being influenced by the beliefs themselves, were affected by the experimental variables. Further analysis revealed three tendencies responsible for the effects (two-tailed *t* tests):

1. For the contextual variable, the ascription to "intelligent, respectable people" slightly increased the total number of "correct" recalls, especially in comparison with the approach of "What do *you* think will happen when . . .?" ($p < .05$; other comparisons approached but did not reach .05).

2. Also for the variable of context, the ascription to "ignorant, superstitious" people produced less total agreement, except in comparison with "some people" ($p < .05$); but the other three ascriptions did not differ significantly from one another.

3. For the inquiry variable, the request to supply reasons increased slightly but significantly the total number of "don't know" attitudes ($p < .05$), even when a set of beliefs did not have to be so justified but followed a set which did. This variable, moreover, significantly increased the reaction times to the questions eliciting both knowledge ($p < .01$) and attitude ($p < .05$), again both for beliefs in and also not in the sets to which the question "why" was affixed.

Aside from the association with assent and one other still to be mentioned, no other correlate of availability of knowledge could be found in the data. The measure of assent, on the other hand, was significantly related to other variables in both studies. Among the adults, those agreeing with relatively few beliefs tended to be different in innumberable respects from those agreeing with relatively many beliefs; for example (all p's $< .01$, two-tailed tests):

1. They had attended school longer.
2. They had skins whose color was judged by the investigator to be on the average appreciably lighter.
3. They were less discontent—based upon answers to 15 questions that were scaled in an a priori manner.
4. They scored higher on one or more of three tests involving the ability to recognize geometrical forms, to discover embedded figures, and to sort figures according to some abstract principles.
5. They could verbalize more accurately the principle that they had used to sort the figures in the last mentioned test.
6. They had a stronger tendency to organize their impressions of stereoscopic pictures rather than to report discrete and unrelated details.

In none of the above respects did those scoring high in availability of knowledge differ significantly from those scoring low. School authorities divide the children into two "streams" (sections) on the basis of educational achievement. The brighter group had a mean of 5.1 agreements out a possible 20, the duller one of 8.3 ($p < .01$). The mean number of "correctly" stated beliefs for the corresponding groups was 5.9 and 5.7. The multiple-choice form of the Benton Visual Retention Test (1950), which was given the children as a respite during the interview, was negatively related to attitude agreement ($p < .01$), but was unrelated to availability of knowledge. Without exception, the differences between those agreeing with few and with many traditional beliefs indicate that those departing from traditions had a higher social status, better education, or greater ability.

DISCUSSION

A low but positive association between the availability of knowledge concerning traditional beliefs and agreement with those beliefs has been demonstrated. That association, however, may have resulted not from an interaction in the past (as has been here presumed) but from a methodological artifact: since those not recalling the "correct" belief were in effect being gently corrected by the investigator, they may have been too distressed to indicate their assent or otherwise have not been motivated to do so. The continued association between availability and assent even during the manipulation of the variables of context and inquiry argues indirectly against such a situational explanation.

In addition, more direct evidence indicates that, among comparable informants, virtually identical attitudes were expressed when the questions pertaining to availability of knowledge were *not* asked. Before those questions were included in the interviewing schedule for the adults, 47 people had expressed only their attitude on 5 or 6 of the 10 beliefs eventually presented to the 65 persons of the second study reported here. When a comparison was made between adults of similar status (rural, lower-class) who expressed attitude alone and those who expressed attitude after recalling knowledge, no significant differences appeared. The 12 workers in St. Ann were interviewed twice: in the first session only attitude was ascertained; in the second, about 10 weeks later, both knowledge and attitude were elicited. For the five beliefs thus tested twice, no significant differences in attitude emerged. In an experiment conducted among Jamaican high school girls for a different purpose and hence under relatively different conditions (Doob, 1958), six of the same beliefs used with the children in the present study were ascertained: in two instances, the results were statistically identical and in four the older, better educated girls showed significantly *less* agreement.

The magnitude of the range of "correct" recalls for the various beliefs in the two studies—from 11 to 83 %—suggests that more than attitude affected availability of knowledge. Doubtless wording of the question that tested availability played a role. Additional clues in the question, for example, significantly raised the percentage of "correct" recalls among the children in comparison with the adults whenever those clues were provided as the questions were adapted to the children; when there were no changes, the two groups tended to be very similar. Other differences must have resulted from experiences that were not explored in the investigation. In fact, the urban dwellers in the adult sample could "correctly" recall more beliefs than those in the rural areas ($p < .02$), even though the former assented to fewer beliefs than the latter ($p < .05$).

Then an incidental finding suggests that the intervening experiences affected the availability of knowledge without necessarily changing attitude. During the long interview with the adults, half were exposed to a brief communication that advocated a belief contrary to a traditional Jamaican attitude toward the relation of food and nutrition. A before-and-after measure in the same interview and a comparison with the other half of the *S*s who served as controls indicated that the communication did not change attitude. Of those whose attitudes had been initially untraditional, however, all six *S*s in the control group could subsequently recall the "correct" traditional belief; whereas only 4 of the 10 *S*s in the experimental group who were exposed to the communication's arguments against the traditional view could produce "correct" recall. No such differences appeared when the comparison involved *S*s whose initial attitudes were traditional. Virtually all *S*s could recall part of the communication itself when requested still later in the interview to do so. The numbers are too small to permit a conclusion to be drawn, but the results support the view that the availability of knowledge concerning traditional beliefs is affected not only by attitude but also by attitude in conjunction with competing beliefs. The association between misfortune and 13 may occur less frequently when the validity of the belief is doubted *and* when misfortune is attributed to other causes.

The relatively few effects of the contextual and inquiry variables suggest that during the interview people's more enduring predispositions were not appreciably influenced by the situational

factors here manipulated. Among the children, only a high prestige frame of reference slightly increased the number of "correctly" recalled beliefs, and only the ascription of a belief to a low-prestige group affected the expression of attitude toward that belief. When informants were asked to supply reasons for their beliefs, they became more cautious and were slower in expressing their attitudes, but if they were willing to express themselves after some delay and as has been established among American college students (Carper & Doob, 1953–54), their expressed attitude did not seem to be affected. From a practical standpoint, such stability bolsters confidence in the reliability of the interviewing procedure, a point of some importance today as psychological procedures are increasingly employed in so-called underdeveloped areas.

SUMMARY

The availability of knowledge concerning a traditional belief was determined by asking 65 adults and 93 school children in Jamaica, B.W.I., to supply that belief in response to an open-ended question, assent by having them signify their agreement or disagreement. A significantly positive relation between availability of knowledge and assent was demonstrated. That relation persisted even when the form of questioning was experimentally manipulated. Various attributes of higher status in Jamaica and greater ability on performance tests were found to be associated with assent but not with availability of knowledge.

REFERENCES

BENTON, A. L. A multiple choice type of the Visual Retention Test. *AMA Arch. Neurol. Psychiat.*, 1950, **64**, 699–707.

CARPER, J., & DOOB, L. W. Intervening responses between questions and answers in attitude surveys. *Publ. opin. Quart.*, 1953–54, **17**, 511–519.

CARR, MADELINE. *Personality and conflict in Jamaica.* Liverpool: Liverpool Univer. Press, 1952.

DOOB, L. W. An introduction to the psychology of acculturation. *J. soc. Psychol.*, 1957, **45**, 143–160.

DOOB, L. W. The effect of the Jamaican patois on attitude and recall. *Amer. Anthrop.*, 1958, **60**, 574–575.

FESTINGER, L. *A theory of cognitive dissonance.* Evanston: Row, Peterson, 1957.

HARDING, J., KUTNER, B., PROSHANSKY, H., & CHEIN, I. Prejudice and ethnic relations. In G. Lindzey (Ed.), *Handbook of social psychology.* Cambridge: Addison-Wesley, 1954. Pp. 1021–1061.

ROSENBERG, M. J. Cognitive structure and attitudinal affect. *J. abnorm. soc. Psychol.*, 1956, **53**, 367–372.

TAYLOR, C. S. Color and class: A comparative study of Jamaican status groups. Unpublished doctoral dissertation, Yale Univer., 1955.

Received June 18, 1958.

SUBJECTIVE PROBABILITIES OF INTERPERSONAL RELATIONSHIPS[1]

CLINTON B. DE SOTO AND JAMES L. KUETHE

Johns Hopkins University

OFTEN a person is uncertain whether or not some interpersonal relationship exists. He may not know whether or not Jim likes Bill. But if called upon, he can ordinarily form a subjective probability that the relationship exists, taking into account whatever relevant information he has. For a given pair of people a and b, and a given relationship R, such as "likes," he may form a subjective probability that a likes b, $P(aRb)$, which may take on many values. Such subjective probabilities are of psychological interest both as factors in people's decisions and actions with regard to interpersonal relationships and as manifestations or derivatives of their "naive psychology" in Heider's sense (1958).

This article is concerned with $P(aRb)$ for various relationships under a condition where the subject (S) is given minimal information about a and b and under conditions where some information is provided regarding other participations of a or b in the given relationship. An attempt is made to answer questions like these: Which relationships have the highest and which the lowest subjective probabilities? Which are ascribed to general traits of the participants? Which are seen as tending to be symmetric? Which asymmetric? Which transitive?

[1] This work was supported in part by Grant NSF-G4827 from the National Science Foundation.

TABLE 1
QUESTIONS USED TO OBTAIN SUBJECTIVE PROBALITIES FOR THE RELATION "LIKES"

Question	Probability
1. Does Les like Al?	$P(aRb)$
2. Cal likes Harvey. Does Cal like Ralph?	$P(aRb \mid aRc)$
3. Jack likes Dave. Does Russ like Dave?	$P(aRb \mid cRb)$
4. Ernie likes Len. Does Len like Ernie?	$P(aRb \mid bRa)$
5. Bill likes Norman. Norman likes Glenn. Does Bill like Glenn?	$P(aRb \mid aRc, cRb)$

METHOD

There have been many attempts to systematize the myriad interpersonal relationships that are named in English. For the present study, ten varied, widely representative relationships were selected, without reliance on any one of the formal classifications. The relationships selected are "likes," "trusts," "feels superior to," "is happier than," "confides in," "dominates," "lies to," "dislikes," "is afraid of," and "hates."

For each relationship, $P(aRb)$ was obtained for the condition of minimal information and for conditions in which information was given that aRc holds, that cRb holds, that bRa holds, or that aRc and cRb hold. For the remainder of this article, the probabilities for the latter conditions will be expressed as conditional probabilities thus: $P(aRb \mid aRc)$, $P(aRb \mid cRb)$, $P(aRb \mid bRa)$, $P(aRb \mid aRc, cRb)$. $P(aRb)$ will be reserved for the condition of minimal information.

To determine the subjective probabilities held by *S*s, questions were prepared of the sort shown in Table 1.

E told the *S*s that people in general are pretty good at making guesses and predictions about the behavior of other people and relationships between them. He said the purpose of the experiment was to learn something about how they do it. He told them that they were to answer questions about the likelihood of various relationships between people under various circumstances. The exact names of the people referred to in the questions were not important—they could be anybody—but a repetition of a name in the same question referred to the same person. The *S*s were to assume that all people mentioned in a given question were acquainted. The *S*s answered each question by circling one of the following answers, which were intended to represent a seven-level scale of probabilities: definitely no, very probably no, probably no, chances about even, probably yes, very probably yes, definitely yes.

Two forms of the questionnaire were prepared which used different names for many corresponding questions (male names were used throughout) and which presented the relationships in different orders. It was found that these differences had no effect on the *S*s' answers.

The questionnaire was given to 100 male undergraduates of the Johns Hopkins University. Their answers were converted to numerical probabilities by giving the seven levels values of 0, 1/6, 1/3, 1/2, 2/3, 5/6, and 1. This is an arbitrary mapping, but it gives results that are similar in all important respects to those that would be obtained from most other plausible mappings.

RESULTS

In Table 2 are shown the mean subjective probabilities of aRb for the various relationships and conditions. The relationships are ordered from top to bottom by decreasing size of $P(aRb)$, the probability of aRb under conditions of minimal information.

$P(aRb)$ for the various relationships may be regarded as a baseline probability. This baseline probability is highest for "likes" and "trusts," which are positive affective or evaluative reactions

TABLE 2
SUBJECTIVE PROBABILITIES OF INTERPERSONAL RELATIONSHIPS

Relationship	Probability				
	$P(aRb)$	$P(aRb \mid aRc)$	$P(aRb \mid cRb)$	$P(aRb \mid bRa)$	$P(aRb \mid aRc, cRb)$
likes	.59	.58	.60	.74	.66
trusts	.59	.58	.63	.73	.66
feels superior to	.57	.60	.56	.37	.71
is happier than	.49	.54	.55	.09	.89
confides in	.48	.49	.53	.73	.53
dominates	.45	.57	.55	.12	.70
lies to	.44	.63	.45	.60	.57
dislikes	.42	.49	.52	.68	.45
is afraid of	.41	.51	.51	.28	.70
hates	.35	.47	.46	.67	.42

Note.—A difference between probabilities of .04 or greater is significant at the .01 level by two-tail t test.

of *a* to *b*. The lowest baseline probability, and the one farthest from .50, was obtained for "hates," an extremely negative reaction of *a* to *b*. Two other negative relationships, "dislikes" and "is afraid of," obtained the next lowest baseline probabilities. The remaining relationships have probabilities intermediate to those of the more clearly positive and negative ones.

The difference between $P(aRb \mid aRc)$ and $P(aRb)$ reflects the extent to which the relationship is attributable to a general trait of *a*. This difference is largest for "lies to"—if *a* lies to *c*, he is expected to lie to *b*. Lying is seen as very much attributable to a general trait of the liar. For most of the other relationships, $P(aRb \mid aRc)$ is greater than $P(aRb)$ in some degree, indicating that the relationships are attributed in some degree to general traits of *a*. The difference is negligible for "likes," "trusts," and "confides in." The fact that *a* likes, trusts, or confides in *c* does not make him any more or less likely to like, trust, or confide in *b*. People are expected not to be general likers, trusters, or confiders, but at the same time they are expected not to limit their liking, trusting, and confiding to one person.

The difference between $P(aRb \mid cRb)$ and $P(aRb)$ reflects the extent to which the relationship is attributed to a general trait of *b*. This difference is significant for several of the relationships, especially the negative affective ones, but it is not extremely large for any of them.

$P(aRb \mid bRa)$ is a measure of the symmetry of *R*. If this probability is 1.00, *R* is perfectly symmetric; if it is 0.00, *R* is perfectly asymmetric (De Soto & Kuethe, 1958). When it is between 1.00 and 0.00, but above $P(aRb)$, *R* is seen as tending to be symmetric; when it is between 1.00 and 0.00, but below $P(aRb)$, *R* is seen as tending to be asymmetric. In the case of behavioral or affective relationships, $P(aRb \mid bRa)$ is the probability that *a*'s behavior or affect is reciprocated by *b*.

$P(aRb \mid bRa)$ is substantially above or below $P(aRb)$ for all the relationships considered here. For the positive relationships, "likes," "trusts," and "confides in," the previously discussed conditional probabilities were not greatly different from $P(aRb)$, but these relationships are all seen as strikingly symmetric. People are expected to like, trust, and confide in those who like, trust, and confide in them. The negative relationships, "lies to," "dislikes," and "hates," are also seen as strongly symmetric—people are expected to lie to, dislike, and hate those who lie to, dislike, and hate them. In general, the lower $P(aRb)$ for a relationship, the greater its tendency to be symmetric.

Four relationships were perceived as relatively asymmetric: "feels superior to," "is happier than," "dominates," and "is afraid of." The relationships "is happier than" and "dominates" are objectively asymmetric, and the values of $P(aRb \mid bRa)$ obtained for them, although not 0.00, are by far the lowest probabilities in Table 1.

$P(aRb \mid aRc, cRb)$, when compared with the other probabilities, is a measure of the transitivity of *R*. The relationships which were perceived as most strongly transitive are the four which were perceived as asymmetric, "feels superior to," "is happier than," "dominates," and "is afraid of." This combination of the asymmetric and transitive properties indicates that these relationships are seen as partial or complete orderings (De Soto & Kuethe, 1958). "Likes" and "trusts" are seen as fairly strongly transitive, "confides in" as rather weakly so. For "lies to," "dislikes," and "hates," $P(aRb \mid aRc, cRb)$ is larger than $P(aRb)$ but less than $P(aRb \mid aRc)$. If *a* hates *c*, he is a hater and is quite likely to hate *b*—unless *c* hates *b*, in which case *a* has reason to be sympathetic toward *b*. These relationships cannot be said to be either transitive or intransitive.

DISCUSSION

Like Homans (1950), the *S*s assigned high probabilities to positive affective relationships. The premise that positive affective relationships are customary implies that knowledge of their occurrence gives rather little information about the participants. Thus the *S*s were right in not drawing inferences about general traits of the participants in these relationships, even though they did draw inferences about general traits of the participants in rarer relationships.

Their perception of many of the relationships as tending strongly to be symmetric is in keeping with Heider's (1958) discussion of "balance," with Newcomb's (1953) discussion of "strain toward symmetry," with Malinowski's (1932) insistence on the importance of reciprocal and symmetric interpersonal relationships in all societies, and with Freud's (1955) claim that the law of talion is deeply rooted in human feelings. It is important, however, to distinguish the relationships that were perceived as symmetric and transitive, "likes" and "trusts," from those that were perceived as symmetric but not particularly transitive, "confides in," "lies to," "dislikes," and "hates." The theory of structural balance (Heider, 1958; Cartwright & Harary, 1956) has some implications for the perceived transitivity of symmetric relationships. It implies that the positive relationships should be seen as transitive, which is

the case for "likes" and "trusts" but not for "confides in." Perhaps "confides in" should not be subsumed under the class of positive relationships treated by the theory. Heider's version of the theory has no clear implication about the transitivity of negative relationships, but the version by Cartwright and Harary implies that they should be seen as intransitive, an implication that receives slight support in the data.

An alternative view of these results is suggested by the data of De Soto and Kuethe (1958). They found that 9- and 10-year-old children were much more likely to perceive "likes" as intransitive than as transitive, a clear violation of the balance theory. It can be inferred that a relationship which is perceived as symmetric and intransitive or nontransitive is perceived as essentially a pair-wise interchange, while one perceived as symmetric and transitive is perceived as dividing people into equivalence classes or groups (De Soto & Kuethe, 1958). Perhaps the fundamental primitive way of regarding interpersonal relationships is as pair-wise interchanges, as the children regarded "likes" and as the present sample of college students regarded "confides in," "lies to," "dislikes," and "hates." Perceiving "likes" and "trusts" as transitive, dividing people into cliques or friendship groups, may be an outcome of long experience in observing people. This experience presumably teaches that if one person likes another it most likely means that they are in the same friendship group, not that the one has a general trait of liking or the other a general trait of likeableness, and not that they have an exceptional, pair-wise affinity.

The *Ss*' perception of "is happier than" as asymmetric and transitive, ordering people, is objectively accurate, but it is noteworthy that "dominates," "feels superior to," and "is afraid of" were similarly perceived. The *Ss* evidently regarded these reactions as based on objective differences on a dimension of power or strength, an unnecessary inference that suggests a great readiness on their part to think in terms of power-based orderings. *Ss* other than college students might not show this readiness. De Soto and Kuethe (1958) found that adults who had never attended college, unlike college students, often ascribed the intransitive property to "dominates," thus denying that it orders people.

Some additional findings by De Soto and Kuethe (1958) bear on the present study. They found that when no information is provided about a relationship, noncollege adults are more likely than college students to venture a guess that it exists, in the case of "likes" and "loves," or does not exist in the case of "dislikes." Noncollege adults are also more likely to ascribe the symmetric property to these relationships. Crudely translating these findings into subjective probabilities, it appears that for these relationships the noncollege adults' subjective probabilities are essentially like those of college students, but more extreme. Perhaps a college education induces chariness in making inferences about these relationships. In any case there is substantial evidence that subjective probabilities of interpersonal relationships vary with *Ss*' age, education, and very likely with other traits, and interpretations of subjective probabilities obtained from restricted samples of *Ss* should accordingly be cautious.

SUMMARY

Subjective probabilities of the existence of various interpersonal relationships under various conditions were obtained with a questionnaire given male college students.

Positive affective relationships were regarded as generally most probable, and negative affective relationships as least probable.

The relationships "likes" and "trusts" were perceived as tending to be symmetric and transitive, a finding interpreted as indicating that these relationships were regarded as dividing people into friendship groups or cliques. The relationships "confides in," "lies to," "dislikes," and "hates" were perceived as tending to be symmetric but not particularly transitive, a result interpreted as meaning that these relationships were regarded as essentially pair-wise interchanges. The occurrence of the relationships "lies to," "dislikes," and "hates," was considered to give some information about general traits of the persons involved, increasing the likelihood they would be similarly involved elsewhere, but this was not so for "likes," "trusts," and "confides in."

The relationship "is happier than" was accurately perceived as asymmetric and transitive. The relationships "feels superior to," "dominates," and "is afraid of" were also perceived as tending strongly to be asymmetric and transitive, a finding interpreted as indicating that these *Ss* had a great readiness to think in terms of power-based orderings.

REFERENCES

CARTWRIGHT, D., & HARARY, F. Structural balance: A generalization of Heider's theory. *Psychol. Rev.*, 1956, **63**, 277–293.

DE SOTO, C. B., & KUETHE, J. L. Perception of mathematical properties of interpersonal relationships. *Percept. mot. Skills*, 1958, **8**, 279–286.

FREUD, S. *Totem and taboo.* (Stand. ed., Vol. XIII) London: Hogarth, 1955.

HEIDER, F. *The psychology of interpersonal relations.* New York: Wiley, 1958.

HOMANS, G. C. *The human group.* New York: Harcourt, Brace, 1950.

MALINOWSKI, B. *Crime and custom in savage society.* New York: Harcourt, Brace, 1932.

NEWCOMB, T. M. An approach to the study of communicative acts. *Psychol. Rev.*, 1953, **60**, 393–404.

Received February 2, 1959. Prior publication.

PSYCHOLOGICAL MONOGRAPHS

General and Applied

Collector's items of the future?*

A Dual Approach to Rorschach Validation, by J. O. Palmer. No. 326. *$1.00*

A Prognostic Scale for Shock Therapy, by M. J. Feldman. No. 327. *$1.00*

The Crystallization of Attitudes Toward the Negro as a Function of Education, by L. A. Kahn. No. 330. *$1.50*

Studies in Reputation: I. Sex and Grade Differences in School Children's Evaluations of Their Peers. II. The Diagnosis of Social Adjustment, by R. D. Tuddenham. No. 333. *$1.50*

Personal Security as Related to Station in Life, by A. L. Knutson. No. 336. *$1.00*

The Performance of Schizophrenic and Normal Individuals Following Frustration, by Harold Wilensky. No. 344. *$1.00*

Parents' Attitudes vs. Adolescent Hostility in the Determination of Adolescent Sociopolitical Attitudes, by Kenneth Eisdant. No. 345. *$1.00*

Personality of the Convulsive Patient in Military Service, by T. W. Richards. No. 346. *$1.00*

Judging Interests from Expressive Behavior, by N. L. Gage. No. 350. *$1.00*

The Effects of Analeptic Drugs in Relieving Fatigue, by R. H. Strasmich and A. C. Ivy. No. 365. *$1.00*

(*Random selection from Volumes 65, 66, and 67. Write for complete list 1951–1953)

If any of these monographs are missing from your library—lost, strayed, or stolen—order now, before they go out of print, from

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street N.W.
Washington 6, D. C.

PSYCHOLOGY AND REHABILITATION

Beatrice A. Wright, Ph.D.
Editor

Based on the proceedings of an Institute on the roles of psychology and psychologists in rehabilitation, sponsored by the American Psychological Association and supported by a grant from the Office of Vocational Rehabilitation, Department of Health, Education and Welfare.

The volume examines the principles and assumptions underlying the modern philosophy of rehabilitation, paying special attention to the roles and functions of psychologists as members of an interdisciplinary team. Problems of training are explored in detail and in the context of the changing nature of the field. Areas of needed research are highlighted.

Price $1.50

Order from

American Psychological Association, Inc.
Dept. PAR
1333 Sixteenth Street, N.W.
Washington 6, D. C.